The Discrete Fourier Transform

2.1. Sums of roots of unity

Suppose we are given a polynomial in factored form, say

$$P(z) = \sum_{k=0}^{n} a_k z^k = a_n \prod_{j=1}^{n} (z - z_j) \tag{2.1}$$

with $a_n \neq 0$. When the product is expanded, the coefficient of z^{n-1} is

$$-a_n \sum_{j=1}^{n} z_j \,.$$

Thus the sum of the n roots of P is

$$z_1 + z_2 + \cdots + z_n = -\,\frac{a_{n-1}}{a_n}\,.$$

In the case of the polynomial $P(z) = z^n - 1$, whose factorization is given in (1.19), we have $a_{n-1} = 0$ when $n > 1$, and hence

$$\sum_{j=1}^{n} e^{2\pi i j/n} = 0\,. \tag{2.2}$$

This is not surprising, since the roots $e^{2\pi i j/n}$ are the vertices of a regular n-gon inscribed in the unit circle $|z| = 1$, and it might be expected that their center of gravity would lie at the center. In what follows, we need to understand not just the sum of the roots of unity, but also the sum of the k^{th} powers of these roots, for various k.

Starting now, and continuing through the rest of our work, many formulas will include an expression of the form $e^{2\pi i x}$ where the quantity x may itself involve several terms. To ease our typography we now adopt a notation introduced many decades ago by the great Russian mathematician I. M. Vinogradov:

$$e(x) = e^{2\pi i x}\,. \tag{2.3}$$

We refer to this as *the complex exponential with period one.*

Theorem 2.1. *Let q be a positive integer. Then for any integer k,*

$$\sum_{n=1}^{q} e(kn/q) = \begin{cases} q & (q|k), \\ 0 & (\textit{otherwise}). \end{cases}$$

Proof. If q divides k evenly among the integers (in symbols $q|k$), then k/q is an integer, so $e(kn/q) = 1$ for all n by (1.18), and hence the q terms sum to q. Suppose that q does not divide k (in symbols $q \nmid k$). The terms $e(kn/q)$ form a geometric progression with common ratio $e(k/q) \neq 1$ from one term to the next, so by the formula (0.2) for the sum of a geometric progression we see that

$$\sum_{n=1}^{q} e(kn/q) = \frac{e(k(q+1)/q) - e(k/q)}{e(k/q) - 1} = e(k/q)\frac{e(kq/q) - 1}{e(k/n) - 1}\,.$$

Here $e(kq/q) = e(k) = 1$, so the numerator is 0, and the denominator is nonzero, so the sum is 0 in this case. □

Example 2.1. Let $\zeta = e(1/5)$, and put

$$a = 2\cos 2\pi/5 = \zeta + \zeta^4, \qquad b = 2\cos 4\pi/5 = \zeta^2 + \zeta^3\,.$$

Then $a + b = \zeta + \zeta^2 + \zeta^3 + \zeta^4 = -1$ by (2.2), and $ab = (\zeta + \zeta^4)(\zeta^2 + \zeta^3) = \zeta + \zeta^2 + \zeta^3 + \zeta^4 = -1$. Thus

$$(z - a)(z - b) = z^2 - (a + b)z + ab = z^2 + z - 1\,.$$

Hence by the formula for the roots of a quadratic polynomial, a and b are $(-1 \pm \sqrt{5})/2$ in some order. But clearly $a > 0$ and $b < 0$, so we conclude that

$$\cos\frac{2\pi}{5} = \frac{-1 + \sqrt{5}}{4}, \qquad \cos\frac{4\pi}{5} = -\,\frac{1 + \sqrt{5}}{4}\,. \tag{2.4}$$

Exercises

Let $z_1, z_2, \ldots, z_n$ be given complex numbers. For an integer k, $1 \le k \le n$, multiply k of them together, in all $\binom{n}{k}$ different possible ways, and sum the results. This quantity is denoted σ_k, and is called the k^{th} *elementary symmetric function* of the z_j. By convention, we put $\sigma_0 = 1$.

1. (a) Show that $\cos(\pi - \theta) = -\cos\theta$.
   (b) Deduce that
   $$\cos\frac{\pi}{5} = \frac{1 + \sqrt{5}}{4}\,.$$
2. Let $\zeta = e(1/7)$, and put $a = \zeta + \zeta^6 = 2\cos 2\pi/7$, $b = \zeta^2 + \zeta^5 = 2\cos 4\pi/7$, and $c = \zeta^3 + \zeta^4 = 2\cos 6\pi/7$.
   (a) Show that $a + b + c = -1$.

(b) Show that $ab + bc + ca = -2$.
(c) Show that $abc = 1$.
(d) Conclude that $(z-a)(z-b)(z-c) = z^3 + z^2 - 2z - 1$.

3. Suppose that $P(z)$ is given as in (2.1). Show that $\sigma_k = (-1)^k a_{n-k}/a_n$.
4. Suppose that

$$Q(z) = \prod_{j=1}^{n} (1 - z_j z)\,.$$

(a) Show that

$$Q(z) = 1 - \sigma_1 z + \sigma_2 z^2 - \cdots + (-1)^n \sigma_n z^n\,.$$

(b) Show that

$$Q'(z) = -\sigma_1 + 2\sigma_2 z - \cdots + (-1)^n n \sigma_n z^{n-1}\,.$$

(c) Show that the logarithmic derivative of $1 - z_j z$ is $-z_j/(1 - z_j z)$.
(d) Show that

$$\frac{Q'(z)}{Q(z)} = -\sum_{j=1}^{n} \frac{z_j}{1 - z_j z}\,.$$

(e) Suppose that z is so small that $|z_j z| < 1$. Explain why

$$\frac{1}{1 - z_j z} = \sum_{m=0}^{\infty} (z_j z)^m\,.$$

(f) Let $s_m = z_1^m + z_2^m + \cdots + z_n^m$. This is called a *power sum.* (Guess why.) Deduce that

$$Q'(z) = -Q(z) \sum_{m=0}^{\infty} s_{m+1} z^m$$

when $|z|$ is sufficiently small. By comparing the constant terms on both sides, deduce that $\sigma_1 = s_1$. By comparing the coefficients of z on the two sides above, deduce that $2\sigma_2 = \sigma_1 s_1 - s_2$. In general, by comparing the coefficient of z^{m-1} on each side, deduce that

$$\sum_{k=1}^{m} (-1)^{k-1} \sigma_{m-k} s_k = \begin{cases} m\sigma_m & (1 \le m \le n), \\ 0 & (m > n). \end{cases} \tag{2.5}$$

These are the *Newton–Girard formulæ.* From the theory of linear recurrences (see §F.4) we know that the s_m must satisfy a linear recurrence. By these formulæ the linear recurrence is made explicit. Given the σ_k, we can use these formulæ to compute the s_m. In the reverse direction, if A is an $n \times n$ matrix and the z_j are its eigenvalues, then $\operatorname{tr} A^m = s_m$, so by computing $A^2, A^3, \ldots, A^n$ one can compute the s_m, and hence the σ_k, and consequently the characteristic polynomial of A.

5. We apply the above to $Q(z) = 1 - z^q$.

(a) Explain why in this case

$$s_m = \sum_{n=1}^{q} e(mn/q)\,.$$

(b) Show that $\sigma_k = 0$ for $0 < k < q$, and that $\sigma_q = (-1)^{q-1} q$.

(c) Use the Newton–Girard formulæ to show that $s_m = 0$ for $m = 1, 2, ..., q-1$ and that $s_q = q$.
(d) Observe that in the present situation, $s_{m+q} = s_m$.
(e) Use these results to complete a second proof of Theorem 2.1.

2.2. The Transform

We say that $f(n)$ is an *arithmetic function* if it is defined for all integers. We also say that an arithmetic function is *periodic with period* q if $f(n + mq) = f(n)$ for all integers m. For an arithmetic function f with period q, we define the *Discrete Fourier Transform* (DFT) to be

$$\widehat{f}(k) = \sum_{n=1}^{q} f(n)e(-nk/q) \tag{2.6}$$

for each integer k. The numbers $\widehat{f}(k)$ also have period q, since $e(-n(k + mq)/q) = e(-nk/q)e(-nm) = e(-nk/q)$ for all n. The $\widehat{f}(k)$ are defined in terms of the $f(n)$, but this can be reversed:

Theorem 2.2. *Suppose that $f(n)$ has period q, and let $\widehat{f}(k)$ be defined as in* (2.6). *Then*

$$f(n) = \frac{1}{q}\sum_{k=1}^{q} \widehat{f}(k)e(kn/q)\,. \tag{2.7}$$

This is the Fourier expansion of f, by which we see that any arithmetic function with period q can be expressed as a linear combination of the basic functions $e(kn/q)$. Conventions vary among authors as to where the factor $1/q$ should go. Some writers put q on the right hand side of (2.6) rather than in the denominator in (2.7), while others put a factor $1/\sqrt{q}$ in both formulæ.

Proof. From the definition (2.6), we see that the right hand side above is

$$= \frac{1}{q}\sum_{k=1}^{q} e(kn/q)\sum_{a=1}^{q} f(a)e(-ak/q)\,.$$

In (2.6) we used n to index the sum, but in the above we have used a new letter, a, since n is now in use. By inverting the order of summation we see that the above is

$$= \frac{1}{q}\sum_{a=1}^{q} f(a)\sum_{k=1}^{q} e(k(n-a)/q)\,.$$

By Theorem 2.1, the inner sum is q if $q|(n-a)$. Otherwise the inner sum is 0. For $1 \le a \le q$ there is exactly one a such that $q|(n-a)$, and for this a we have $f(a) = f(n)$ since f has period q. Hence the above is $f(n)$, as desired. □

When $q|(n-a)$, as we encountered above, we say that a is *congruent to* n *modulo* q, and in symbols we write $a \equiv n \pmod{q}$.

The functions $f_k(n) = e(kn/q)$ have period q, and the import of Theorem 2.2 is that they form a basis for the vector space of all functions with period q. When considering vectors in n-dimensional real space $\mathbb{R}^n$, the vectors $\boldsymbol{u} = [u_j]$ and $[\boldsymbol{v}] = [v_j]$ have the inner product (or 'dot product') $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{j=1}^n u_j v_j$. When dealing with n-dimensional vectors with complex elements, this is modified as follows:

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{j=1}^{n} u_j \overline{v_j}\,. \tag{2.8}$$

Thus, while in the real case $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$, in the complex case we have

$$\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}\,. \tag{2.9}$$

With this understanding, we see by Theorem 2.1 that the functions f_k are not merely linearly independent; they are orthogonal. Indeed, in (2.6) and (2.7) we witness the usual computation of the coordinates of a vector with respect to an orthogonal basis.

In (2.7) we have expressed f as a linear combination of the basis vectors f_k. It follows by principles of linear algebra that this is the unique linear combination that yields f, but we give a direct computational proof of this result.

Theorem 2.3. *Suppose that f has period q, and that $c(1), c(2), \ldots, c(q)$ are numbers such that*

$$f(n) = \frac{1}{q}\sum_{k=1}^{q} c(k)e(kn/q) \tag{2.10}$$

for all n. Then $c(k) = \widehat{f}(k)$ for all k.

Proof. On subtracting (2.7) from (2.10) we find that

$$\sum_{k=1}^{q}(c(k) - \widehat{f}(k))e(kn/q) = 0$$

for all n. Let $g(k) = c(k) - \widehat{f}(k)$. From the definition (2.6), with the roles of k and n reversed, we see that $\widehat{g}(n) = 0$ for all n. From (2.7) we know that

$$g(k) = \frac{1}{q}\sum_{n=1} \widehat{g}(n)e(kn/q)\,.$$

Since $\widehat{g}(n) = 0$ for all n, it follows that $g(k) = 0$ for all k. That is, $c(k) = \widehat{f}(k)$. □

If f has period q and a is an integer, then $f(a+1), f(a+2), \ldots, f(a+q)$ are the numbers $f(1), f(2), \ldots, f(q)$ in a permuted order. Hence

$$\sum_{n=a+1}^{a+q} f(n) = \sum_{n=1}^{q} f(n)\,. \tag{2.11}$$

This applies to such sums as the one in (2.6). Since $f(n)$ has period q and $e(-kn/q)$ has period q, it follows that $f(n)e(-kn/q)$ has period q, and hence

$$\widehat{f}(k) = \sum_{n=a+1}^{a+q} f(n)e(-kn/q) \tag{2.12}$$

for any integer a. Similarly, from (2.7) it follows that

$$f(n) = \frac{1}{q} \sum_{k=a+1}^{a+q} \widehat{f}(k)e(kn/q) \tag{2.13}$$

for any a.

From (2.6), by the triangle inequality we see that

$$|\widehat{f}(k)| \le \sum_{n=1}^{q} |f(n)| \tag{2.14}$$

for all k. Similarly, from (2.7) we deduce that

$$|f(n)| \le \frac{1}{q} \sum_{k=1}^{q} |\widehat{f}(k)| \tag{2.15}$$

for all n.

The DFT is linear in the sense that if f and g have period q and $h(n) = \alpha f(n) + \beta g(n)$ for all n, then from (2.6) we see that $\widehat{h}(k) = \alpha\widehat{f}(k) + \beta\widehat{g}(k)$. Further properties of the DFT are described below.

Theorem 2.4. *Suppose that $f(n)$ has period q.*

(a) *If a is an integer and $g(n) = f(n+a)$ for all n, then $\widehat{g}(k) = \widehat{f}(k)e(ka/q)$ for all k.*

(b) *If b is an integer and $h(n) = f(n)e(bn/q)$ for all n, then $\widehat{h}(k) = \widehat{f}(k-b)$ for all k.*

(c) *If $j(n) = f(-n)$ for all n, then $\widehat{j}(k) = \widehat{f}(-k)$ for all k.*

(d) *If $\ell(n) = \overline{f(n)}$ for all n, then $\widehat{\ell}(k) = \overline{\widehat{f}(-k)}$ for all k.*

Proof. (a) From (2.6) we know that

$$\begin{aligned}
\widehat{g}(k) &= \sum_{n=1}^{q} g(n)e(-kn/q) = \sum_{n=1}^{q} f(n+a)e(-kn/q) \\
&= e(ka/q) \sum_{n=1}^{q} f(n+a)e(-k(n+a)/q) \\
&= e(ka/q) \sum_{n=a+1}^{a+q} f(n)e(-kn/q) = e(ka/q)\widehat{f}(k)
\end{aligned}$$

by (2.12).

(b) From (2.6) we know that

$$\widehat{h}(k) = \sum_{n=1}^{q} h(n)e(-kn/q) = \sum_{n=1}^{q} f(n)e((b-k)n/q) = \widehat{f}(k-b)\,.$$

(c) $\widehat{j}(k) = \sum_{n=1}^{q} f(-n)e(-kn/q)$. Put $m = -n$. Then this is

$$\sum_{m=-q}^{-1} f(m)e(mk/q) = \widehat{f}(-k)$$

by (2.12).

(d) We first replace k by $-k$ in (2.6), and then take complex conjugates of both sides to see that

$$\overline{\widehat{f}(-k)} = \sum_{n=1}^{q} \overline{f(n)}e(-kn/q) = \widehat{\ell}(k)\,.$$

□

Example 2.2. Suppose that $2N < q$. Define $f(n)$ with period q by setting

$$f(n) = \begin{cases} 1 & (-N \le n \le N), \\ 0 & (N < n < q - N). \end{cases}$$

Clearly $\widehat{f}(0) = 2N + 1$. Suppose that $q \nmid k$. Then

$$\widehat{f}(k) = \sum_{n=-N}^{N} e(-kn/q) = e(kN/q) \sum_{n=0}^{2N} e(-kn/q).$$

Here we have a geometric series, so by (0.2) the above is

$$\begin{aligned} &= e(kN/q)\frac{1 - e(-k(2N+1)/q)}{1 - e(-k/q)} \\ &= e(kN/q)\frac{e(-\frac{1}{2}k(2N+1)/q)(e(\frac{1}{2}k(2N+1)/q) - e(-\frac{1}{2}k(2N+1)/1))}{e(-\frac{1}{2}k/q)(e(\frac{1}{2}k/q) - e(-\frac{1}{2}k/q))}\,. \end{aligned}$$

But $e(kN/q)e(-\frac{1}{2}k(2N+1)/q) = e(-\frac{1}{2}k/q)$. We divide the numerator and denominator by $2i$ to see that the above is

$$= \frac{\sin \pi k(2N+1)/q}{\sin \pi k/q}\,.$$

Suppose that $f(n)$ and $g(n)$ are arithmetic functions with period q. The *convolution* of f and g is

$$(f * g)(n) = \sum_{a=1}^{q} f(n-a)g(a), \tag{2.16}$$

which is also an arithmetic function with period q. This is the first of several different convolutions that we shall encounter. The precise definition of a convolution depends on the context, but in all cases one forms a product of the two functions at arguments that sum to one fixed value, and then sums (or integrates) over all such pairs of arguments. Some authors would put a $1/q$ on the right hand side in (2.16). Thus one needs to be aware of the conventions being followed when reading from other sources.

Theorem 2.5. *Suppose that $f(n)$ and $g(n)$ are arithmetic functions with period q, and that $h = f * g$. Then*

$$\widehat{h}(k) = \widehat{f}(k)\widehat{g}(k)\,.$$

Proof. From the definition (2.6) we know that

$$\begin{aligned}
\widehat{h}(k) &= \sum_{n=1}^{q} h(n)e(-kn/q) = \sum_{n=1}^{q} e(-kn/q) \sum_{a=1}^{q} f(n-a)g(a) \\
&= \sum_{a=1}^{q} g(a) \sum_{n=1}^{q} f(n-a)e(-kn/q) \\
&= \sum_{a=1}^{q} g(a)e(-ka/q) \sum_{n=1}^{q} f(n-a)e(-k(n-a)/q)\,. \qquad (2.17)
\end{aligned}$$

In the last sum above, as n runs from 1 to q, $n-a$ runs from $1-a$ to $q-a$. Thus by (2.12), this sum is

$$= \sum_{n=1-a}^{q-a} f(n)e(-kn/q) = \sum_{n=1}^{q} f(n)e(-kn/q) = \widehat{f}(k).$$

This is independent of a, so when we insert this value in (2.17) the remaining sum is just $\widehat{g}(k)$, and the proof is complete. □

Example 2.3. Suppose that f has period q, that $2N < q$, and that

$$f(n) = \begin{cases} 1 & (0 \le n < N), \\ 0 & (N \le n < q). \end{cases}$$

By arguing as in Example 2.2, we find that

$$\widehat{f}(k) = \begin{cases} N & (q|k), \\ e(-\frac{1}{2}k(N-1)/q)\dfrac{\sin \pi kN/q}{\sin \pi k/q} & (q \nmid k). \end{cases}$$

Let $g(n) = f(-n)$, and set $h(n) = f * g$. Then

$$h(n) = \begin{cases} N - |n| & (-N < n < N), \\ 0 & (N \le n \le q-N). \end{cases}$$

By Theorem 2.4(d) we know that $\widehat{g}(k) = \widehat{f}(-k)$. Thus by Theorem 2.5 we deduce that

$$\widehat{h}(k) = \begin{cases} N^2 & (q|k), \\ \Big(\dfrac{\sin \pi kN/q}{\sin \pi k/q}\Big)^2 & (q \nmid k)\,. \end{cases}$$

Let f be an arithmetic function with period q. The average of f over a full period is $\widehat{f}(0)/q$. If the numbers $\widehat{(k)}$ are not too large for $0 < k < q$, we may hope that f has approximately the same average in shorter intervals.

Theorem 2.6. *Let f be an arithmetic function with period q, and suppose that $|\widehat{f}(k)| \le M$ for $0 < k < q$. If a and b are integers with $a < b$, then*

$$\Big|\sum_{n=a}^{b-1} f(n) \; - \frac{b-a}{q}\widehat{f}(0)\Big| \le \frac{2}{\pi} M \log\frac{4q}{\pi}\,.$$

Proof. By the Discrete Fourier expansion of f we see that

$$\sum_{n=a}^{b-1} f(n) = \sum_{n=a}^{b-1} \frac{1}{q}\sum_{k=1}^{q}\widehat{f}(k)e(kn/q) = \frac{1}{q}\sum_{k=1}^{q}\widehat{f}(k)\sum_{n=a}^{b-1} e(kn/q)\,.$$

On separating $k = q$ from the other terms and using the formula for the sum of a geometric progression we deduce that

$$\sum_{n=a}^{b-1} f(n) \; - \frac{b-a}{q}\widehat{f}(0) = \frac{1}{q}\sum_{k=1}^{q-1}\widehat{f}(k)\frac{e(kb/q)-e(ka/q)}{e(k/q)-1}\,. \tag{2.18}$$

Here the right hand side has absolute value not exceeding

$$\frac{2M}{q}\sum_{k=1}^{q-1}\frac{1}{|e(k/q)-1|},$$

and $|e(k/q) - 1| = 2\sin \pi k/q$ for $0 < k < q$, so to complete the proof it suffices to show that

$$\sum_{k=1}^{q-1}\csc\frac{\pi k}{q} < \frac{2}{\pi} q \log\frac{4q}{\pi}\,. \tag{2.19}$$

Since $\csc \pi u$ is convex, it follows that

$$\csc \pi k/q \le q\int_{\frac{k-1/2}{q}}^{\frac{k+1/2}{q}} \csc \pi u\, du\,.$$

On summing this over k we see that the left hand side of (2.19) is

$$\begin{aligned}&\le q\int_{1/(2q)}^{1-1/(2q)} \csc\pi u\, du = \frac{q}{\pi}\Big[\log(\csc\pi u - \cot\pi u)\Big|_{1/(2q)}^{1-1/(2q)} \\ &= \frac{2q}{\pi}\log(\csc\pi/(2q) + \cot\pi/(2q)) = \frac{2q}{\pi}\log(\cot\pi/(4q)) < \frac{2}{\pi}q\log(4q/\pi)\,.\end{aligned}$$

This gives (2.19), so the proof is complete. □

In the notation of the proof just completed, let

$$F(n) = \frac{1}{q}\sum_{k=1}^{q-1}\widehat{f}(k)\frac{e(kn/q)}{e(k/q)-1}\,. \tag{2.20}$$

This acts as a discrete analogue of an antiderivative of $f(n) - \widehat{f}(0)/q$, since the right hand side of (2.18) is $F(b) - F(a)$. The distribution of the values of $F(n)$ is of great interest to us.

Example 2.4. Suppose that S has period q, that $S(0) = 0$, and that $S(n) = 1/2 - n/q$ for $n = 1, 2, \dots, q-1$. This is a discrete sawtooth function. We determine the Discrete Fourier Transform of S. First we note that

$$\widehat{S}(0) = \sum_{n=1}^{q-1} \Big(\frac{1}{2} - \frac{n}{q}\Big) = \frac{q-1}{2} - \frac{1}{q} \cdot \frac{q(q-1)}{2} = 0,$$

which is to be expected, since $S(n)$ is an odd function. Now suppose that $0 < k < q$. We find that

$$\sum_{n=1}^{q-1} ne(-kn/q) = \sum_{n=1}^{q-1} \Big(\sum_{m=1}^{n} 1\Big) e(-kn/q) = \sum_{m=1}^{q-1} \sum_{n=m}^{q-1} e(-kn/q)\,.$$

By the formula (0.2) for the sum of a geometric progression we see that the above is

$$= \sum_{m=1}^{q-1} \frac{e(-km/q) - 1}{1 - e(-k/q)} = \frac{1}{1 - e(-k/q)} \sum_{m=1}^{q-1} \Big(e(-km/q) - 1\Big)\,.$$

By Theorem 2.1 we know that $\sum_{m=1}^{q} e(-km/q) = 0$. In the above, the term $e(-kq/q)$ arising from $m = q$ is missing, and so the sum has value -1. Thus the above is

$$= \frac{1}{1 - e(-k/q)} \big(-1 - (q-1)\big) = \frac{-q}{1 - e(-kd/q)}\,.$$

Hence

$$\widehat{S}(k) = \sum_{n=1}^{q-1} \Big(\frac{1}{2} - \frac{n}{q}\Big) e(-kn/q) = -\frac{1}{2} + \frac{1}{1 - e(-k/q)} = \frac{1 + e(-k/q)}{2(1 - e(-k/q))}\,.$$

We multiply the numerator and denominator by $e(k/(2q))$ to see that this is

$$= \frac{e(k/(2q)) - e(-k/(2q))}{2\big(e(k/(2q)) - e(-k/(2q))\big)} = \frac{\cos \pi k/q}{2i \sin \pi k/q}\,.$$

To summarize, if $S(0) = 0$ and $S(n) = 1/2 - n/q$ for $0 < n < q$, then

$$\widehat{S}(k) = \begin{cases} -\dfrac{i}{2} \cot \pi k/q & (0 < k < q), \\ 0 & (k = 0). \end{cases} \tag{2.21}$$

When we convolve S with a function f we find that if $a < b$, then

$$\frac{1}{2} f(a) + \sum_{a<n<b} f(n) \ + \frac{1}{2} f(b) - \frac{b-a}{q} \widehat{f}(0) = (f * S)(b) - (f * S)(a)\,. \tag{2.22}$$

Thus $(f * S)(n)$ is a close relative of the function $F(n)$ in (2.20), but now the sum of $f(n)$ is symmetrically placed on the interval $[a, b]$, with the first and last terms counted with weight $1/2$.

Example 2.5. Let X be a random variable that takes values $1, 2, \dots, q$ with probabilities $p_1, p_2, \dots, p_q$ respectively. Let Y be a second random variable, independent of X, that takes the values $1, 2, \ldots, q$ with probabilities $r_1, r_2, \ldots, r_q$. Here the p_n and r_n are nonnegative and

$$\sum_{n=1}^{q} p_n = 1, \qquad \sum_{n=1}^{q} r_n = 1\,.$$

We extend the definitions of the sequences p_n and r_n to all integers by making them periodic with period q. We form $X+Y$, and if the result is greater than q, then we subtract q in order to obtain a value in the range $1, 2, \ldots, q$. The probability that $X + Y \equiv n \pmod{q}$ is

$$(p * r)(n) = \sum_{a=1}^{q} p_a r_{n-a}\,.$$

Now suppose that $q_n = 1/q$ for all n. We call a variable Y with this distribution *uniform*. In the above sum, the $1/q$ factors out, and the p_a sum to 1, so we deduce that $(p * q)(n) = 1/q$ for all n. That is, $X + Y$ is uniform $(\text{mod } q)$ if Y is. This has an application to cryptography: Suppose we have a sequence of characters that form our plaintext. These may be thought of as integers in a range $1, 2, \ldots, q$. For example, we could take $q = 26$, and have each integer refer to a letter of the alphabet. More realistically, each character is associated via its ASCII code to a sequence of 8 bits. Thus we have a bit stream, and we work with $q = 2$. We add a quasi-random sequence to this $(\text{mod } q)$ to obtain our cryptotext. The quasi-random sequence is the output of a random-number generator, which starts from a known initial state. If the recipient has an identical random generator and knows the appropriate initial state (the key), then the recipient can generate the same quasi-random sequence and by subtraction recover the original message. If the quasi-random sequence were truly random, then the cryptotext would also be random, and hence the code would be impossible to break. A cryptanalyst seeking to break the code would need to identify and exploit imperfections in the design of the random number generator.

Theorem 2.7. *If $f(n)$ and $g(n)$ are arithmetic functions with period q, then*

$$\sum_{n=1}^{q} f(n)\overline{g(n)} \;=\; \frac{1}{q}\sum_{k=1}^{q} \widehat{f}(k)\overline{\widehat{g}(k)}\,.$$

We refer to the above as "Parseval's Identity for the DFT." The original Parseval Identity pertains to Fourier Series; see Theorems 5.7 and 5.8.

Proof. By the Discrete Fourier Expansion of $g(n)$ found in Theorem 2.2 we see that the left hand side above is

$$\sum_{n=1}^{q} f(n)\overline{\frac{1}{q}\sum_{k=1}^{q}\widehat{g}(k)e(kn/q)} = \frac{1}{q}\sum_{n=1}^{q} f(n)\sum_{k=1}^{q}\overline{\widehat{g}(k)}e(-kn/q)\,.$$

On reversing the order of summation we see that this is

$$= \frac{1}{q}\sum_{k=1}^{q}\overline{\widehat{g}(k)}\sum_{n=1}^{q} f(n)e(-kn/q) = \frac{1}{q}\sum_{k=1}^{q}\widehat{f}(k)\overline{\widehat{g}(k)},$$

as desired. □

The case $g = f$ of this is of special interest.

Corollary 2.8. *If f is an arithmetic function with period q, then*

$$q\sum_{n=1}^{q} |f(n)|^2 \;=\; \sum_{k=1}^{q} |\widehat{f}(k)|^2\,.$$

Example 2.6. Suppose we have a random number generator that produces the integers $1, 2, \ldots, q$ with probabilities $p_1, p_2, \ldots, p_q$. If the random number generator were perfect, then we would have $p_n = 1/q$ for all n. However, the "quasi random" generator is imperfect, so the p_n are only close to $1/q$, say $p_n = (1 + \delta_n)/q$ where the δ_n are small. One prescription for improving such a quasi random generator is to take two samples of it, X and Y, add them, and if the sum is greater than q, then subtract q so that the result lies between 1 and q. The question is to what extent we should believe that this operation gives rise to a distribution that is closer to uniform. Since $\sum_{n=1}^{q} p_n = 1$, it follows that $\sum_{n=1}^{q} \delta_n = 0$. The distribution of $X + Y$ modulo q is

$$\begin{aligned}(p * p)(n) &= \frac{1}{q^2}\sum_{a=1}^{q}(1+\delta_a)(1+\delta_{n-a})\\ &= \frac{1}{q} + \frac{1}{q^2}\sum_{a=1}^{q}\delta_a\delta_{n-a}\\ &= \frac{1}{q}\Big(1 + \frac{1}{q}\sum_{a=1}^{q}\delta_a\delta_{n-a}\Big)\,.\end{aligned}$$

By Theorems 2.2 and 2.5 we see that

$$\sum_{a=1}^{q}\delta_a\delta_{n-a} = \frac{1}{q}\sum_{k=1}^{q}\widehat{\delta}(k)^2 e(kn/q)\,.$$

By the triangle inequality, the left hand side above has absolute value not exceeding

$$\frac{1}{q}\sum_{k=1}^{q}\big|\widehat{\delta}(k)\big|^2 = \sum_{n=1}^{q}|\delta_n|^2$$

by Corollary 2.8. Thus the original relative error δ_n is replaced by a new relative error that does not exceed

$$\frac{1}{q}\sum_{n=1}^{q}|\delta_n|^2\,.$$

In particular, if δ is a number such that $|\delta_n| \le \delta$ for all n, then the new relative error does not exceed δ^2. It may actually be even smaller, since in applying the triangle inequality we may be throwing away some cancelation.

Theorem 2.9. *Let p be an odd prime number, and put*

$$S_{k,p} = \sum_{n=1}^{p} e(kn^2/p)\,.$$

If $p|k$, then $S_{k,p} = p$, but if $p \nmid k$, then $|S_{k,p}| = \sqrt{p}$.

Proof. The first assertion is obvious. Suppose the $p \nmid k$. Then

$$\Big|\sum_{n=1}^{q} e(kn^2/p)\Big|^2 = \sum_{m=1}^{p}\sum_{n=1}^{p} e((m^2-n^2)/p)\,.$$

The change of variable $m = n + h$ gives $m^2 - n^2 = 2nh + h^2$, so the above is

$$= \sum_{h=1}^{p} e(kh^2/p)\sum_{n=1}^{p} e(2khn/p)\,.$$

Now $(2k, p) = 1$, so $p|2kh$ if and only if $p|h$. Thus the above sum over n vanishes unless $h = p$, which gives the result. □

Example 2.7. For an odd prime p, let N_p denote the number of solutions of the congruence $x^2+y^2+z^2 \equiv 0 \pmod p$. That is, among the triples of integers (x, y, z) with $1 \le x, y, z \le p$, the number N_p counts those triples for which $p|(x^2 + y^2 + z^2)$. Since we start with p^3 triples, and any one of them has a chance $1/p$ of producing a quantity that is divisible by p, we anticipate that N_p should be approximately p^2. We show that this is so, in the following precise sense: $|N_p - p^2| < p^{3/2}$. To see this, observe that

$$\sum_{k=1}^{p} e(k(x^2+y^2+z^2)/p) = \begin{cases} p & (p|(x^2+y^2+z^2)),\\ 0 & \text{(otherwise)}.\end{cases}$$

Thus

$$\begin{aligned} pN_p &= \sum_{x,y,z}\sum_{k=1}^{p} e(k(x^2+y^2+z^2)/p)\\ &= \sum_{k=1}^{p}\Big(\sum_{x=1}^{p} e(kx^2/p)\Big)\Big(\sum_{y=1}^{p} e(ky^2/p)\Big)\Big(\sum_{z=1}^{p} e(kz^2/p)\Big) = \sum_{k=1}^{p} S_{k,p}^3 \end{aligned}$$

in the notation of Theorem 2.9. The contribution of $k = p$ on the right hand side is p^3. We subtract this from both sides to see that

$$|pN_p - p^3| = \Big|\sum_{k=1}^{p-1} S_{k,p}^3\Big| \le \sum_{k=1}^{p-1} |S_{k,p}|^3 < p^{5/2} \tag{2.23}$$

since $|S_{k,p}| = \sqrt{p}$ when $p \nmid k$. Thus we have the stated estimate.

Exercises

1. Suppose that f has period q. Use Theorem 2.4(c) to show the following:
   (a) If $f(n) = f(-n)$ for all n, then $\widehat{f}(k) = \widehat{f}(-k)$ for all k.
   (b) If $\widehat{f}(k) = \widehat{f}(-k)$ for all k, then $f(n) = f(-n)$ for all n.
2. Suppose that f has period q, and again use Theorem 2.4(c) to show that $f(n) = -f(-n)$ for all n if and only if $\widehat{f}(k) = -\widehat{f}(-k)$ for all k. (As in the preceding exercise, this requires two arguments.)
3. Suppose that f has period q.

(a) Use Theorem 2.4(d) to show that $\overline{f(n)} = f(n)$ for all n (i.e., f is real-valued) if and only if $\widehat{f}(-k) = \overline{\widehat{f}(k)}$ for all k.
(b) Show that if f is real-valued, then $\widehat{f}(k)e(kn/q) + \widehat{f}(-k)e(-kn/q)$ is real for all k.
(c) Suppose that f is real-valued and that q is even. When $k = q/2$, the two terms above are equal, and therefore the common value must be real. Show that $\widehat{f}(q/2)$ is real, and that $e((q/2)n/q)$ is real.

4. Suppose that f is an arithmetic function with period q. Put $g(k) = \widehat{f}(k)$. Then g has period q. Show that $\widehat{g}(n) = qf(-n)$.

5. Suppose that f is an arithmetic function with period q, and put $g(n) = f(n) + \widehat{f}(-n)/q$. Show that $\widehat{g}(k) = g(k)$.

6. For an arithmetic function with period q, let $\|f\|_1 = \sum_{n=1}^{q} |f(n)|$.
(a) Show that if f and g are arithmetic functions with period q, then $\|f+g\|_1 \le \|f\|_1 + \|g\|_1$.
(b) Show that if f and g are arithmetic functions with period q, then $\|f*g\|_1 \le \|f\|_1\|g\|_1$.

7. Let $f(n)$ be an arithmetic function with period 3, with $f(0) = 0$, $f(1) = 1$, and $f(2) = -1$.
(a) Find $\widehat{f}(k)$ for $k = 0, 1, 2$, and use the fact that $e(\pm 1/3) = -1/2 \pm i\sqrt{3}/2$ to simplify your answers.
(b) There is a constant c such that $\widehat{f}(k) = cf(k)$ for all k. What is this c?
(c) Write $f(n)$ in its discrete Fourier expansion, so that it is expressed as a linear combination of the three basic functions 1, $e(n/3)$, and $e(2n/3)$.

8. Let $f(n)$ be an arithmetic function with period 4, with $f(0) = f(2) = 0$, $f(1) = 1$, and $f(3) = -1$.
(a) Find $\widehat{f}(k)$ for $k = 0, 1, 2, 3$.
(b) There is a constant c such that $\widehat{f}(k) = cf(k)$ for all k. What is this c?
(c) Write $f(n)$ in its discrete Fourier expansion, so that it is expressed as a linear combination of the four functions 1, i^n, $(-1)^n$, $(-i)^n$.

9. Suppose that f is an arithmetic function with period q, that $|f(n)| = 1$ for N values of n in a period, that otherwise $f(n) = 0$, and that $\widehat{f}(k)$ is nonzero for exactly N values of k in a period. Show that $\max_k |\widehat{f}(k)| \ge \sqrt{q}$.

10. Suppose that δ is an arithmetic function with period q. Show that the following two assertions are equivalent:
(i) $\delta(0) = 1$ and $\delta(n) = 0$ for $0 < n < q$;
(ii) $\widehat{\delta}(k) = 1$ for all k.
The function $\delta(n)$ as in (i) above is the *DFT Dirac delta function.*

11. Suppose that f is an arithmetic function with period q. Show that the following two assertions are equivalent:
(i) There is a constant c such that $f(n) = c$ for all n;
(ii) $\widehat{f}(k) = 0$ for all k in the interval $0 < k < q$.

12. Let X and Y be independent random variables as in Example 2.5 with $q = 2$. Show that if $X+Y \pmod 2$ is uniform, then at least one of X or Y is uniform.

13. Let X and Y be independent random variables as in Example 2.5 with $q = 4$, and $p_1 = 1/8$, $p_2 = 3/8$, $p_3 = 3/8$, $p_4 = 1/8$, $r_1 = 1/8$, $r_2 = 3/8$, $r_3 = 1/8$, $r_4 = 3/8$.
    (a) Show that $\widehat{p}(2) = 0$.
    (b) Show that $\widehat{r}(1) = \widehat{r}(3) = 0$.
    (c) Show that $\widehat{p}(0) = \widehat{r}(0) = 1$.
    (d) Deduce that $(p*q)(n) = 1/4$ for all n. That is, $X+Y \pmod 4$ is uniform, even though neither X nor Y are. (Hint: Use Exercise 11.)
14. Let f be an arithmetic function with period q. In the context of Exercise 11, if f is not quite constant, we might measure the extent to which it deviates from its mean $\mu = \widehat{f}(0)/q$, in terms of the sizes of the $\widehat{f}(k)$.
    (a) Let $g(n) = \mu$ for all n. Explain why $\widehat{g}(k) = 0$ for $0 < k < q$.
    (b) Let $h(n) = f(n) - g(n)$. Show that $\widehat{h}(0) = 0$ and that $\widehat{h}(k) = \widehat{f}(k)$ for $0 < k < q$.
    (c) Show that
$$\sum_{n=1}^{q} |f(n) - \mu|^2 = \frac{1}{q} \sum_{k=1}^{q-1} |\widehat{f}(k)|^2 .$$
15. Let $c_1, c_2, \ldots, c_N$ be real or complex numbers, and put
$$S(x) = \sum_{n=1}^{N} c_n e(nx) .$$
    Set $Z = S(0)$ and put
$$Z(q,k) = \sum_{\substack{n=1 \\ n \equiv k \pmod q}}^{N} a_n .$$
    (a) Show that
$$S(a/q) = \sum_{k=1}^{q} Z(q,k) e(ka/q)$$
    for $a = 1, 2, \ldots, q$.
    (b) Show that the average of the numbers $Z(q,k)$ is Z/q.
    (c) Use the preceding exercise to show that
$$\sum_{k=1}^{q} |Z(q,k) - Z/q|^2 = \frac{1}{q} \sum_{a=1}^{q-1} |S(a/q)|^2 .$$
    Let $\mathcal{N}$ be a collection of Z integers in the interval $[1, N]$, and put $a_n = 1$ if $n \in \mathcal{N}$, $a_n = 0$ otherwise. Then $Z(q,k)$ counts the members of $\mathcal{N}$ that are congruent to k modulo q. From the above we see that the members of $\mathcal{N}$ are evenly distributed into residue classes modulo q (in the mean square sense) if the sum on the right is small.
16. Use Example 2.4 and Corollary 2.8 to show that
$$\sum_{k=1}^{q-1} \cot^2 \frac{\pi k}{q} = \frac{(q-1)(q-2)}{3} .$$

(This was established in Exercise 4 by a different method.) Note: One of the formulas in (9.27) may prove to be useful.

17. Suppose that f and g are arithmetic functions with period q. Show that

$$\widehat{fg}(k) = \frac{1}{q}(\widehat{f} * \widehat{g})(k) .$$

18. Suppose that f has period q. Show that

$$\sum_{k=1}^{q} e(hk/q)|\widehat{f}(k)|^2 = q\sum_{a=1}^{q} f(a)\overline{f(a-h)} .$$

19. Let $g(x) = ax^2 + bx + c$ where a, b, c are integers. Suppose that p is an odd prime and that $p \nmid a$. Put

$$S_k(g) = \sum_{n=1}^{p} e(kg(n)/p) .$$

Show that $S_0(g) = p$ and that $|S_k(g)| = \sqrt{p}$ if $p \nmid k$.

20. Let $f(x) = ax^3 + bx^2 + cx + d$ be a cubic polynomial with integer coefficients. Let p be an odd prime, and suppose that $p \nmid a$. Let $S_k(f)$ be defined in the manner of the preceding exercise.
    (a) Let $g_h(x) = f(x+h) - f(x)$. Show that $g_h(x) = 3ahx^2 + (3ah^2 + 2bh)x + (ah^3 + bh^2 + ch)$.
    (b) Show that $|S_k(f)|^2 = \sum_{h=1}^{p} S_k(g_h)$.
    (c) Deduce that $|S_k(f)| \le 2p^{3/4}$ if $p \nmid k$.

2.3. The Fast Fourier Transform

Suppose that f is an arithmetic function with period q. The computation of $\widehat{f}(k)$ involves q multiplications and $q-1$ additions. If we do this for all q values of k, then we perform $\asymp q^2$ arithmetic operations. However, there is a faster way. We confine ourselves to the Fast Fourier Transform (FFT) as it was originally invented when q is a power of 2, but there are other FFTs, including some that work for general q, even prime q.

Theorem 2.10. *Let f have period q, and suppose that all q^{th} roots of unity have already been computed. Let $N(q)$ denote the least number of arithmetic operations required to compute $\widehat{f}(k)$ for all k $(1 \le k \le q)$. Then $N(2q) \le 2N(q) + 3q$.*

Proof. Suppose that f has period $2q$. Put $g(n) = f(2n)$ and $h(n) = f(2n+1)$. Thus g and h have period q, and

$$\begin{aligned}\widehat{f}(k) &= \sum_{n=1}^{2q} f(n)e(-kn/(2q)) \\ &= \sum_{n=1}^{q} f(2n)e(-kn/q) + \sum_{n=0}^{q-1} f(2n+1)e(-k(2n+1)/(2q)) \\ &= \sum_{n=1}^{q} g(n)e(-kn/q) + e(-k/(2q)) \sum_{n=0}^{q-1} h(n)e(-kn/q) \\ &= \widehat{g}(k) + e(-k/(2q))\widehat{h}(k)\,.\end{aligned}$$

We compute $\widehat{g}(k)$ and $\widehat{h}(k)$ for all q values of k. This requires at most $2N(q)$ operations. The computation of $\widehat{f}(k)$ then requires two additional arithmetic operations. When this is done for all $2q$ values of k, we perform $4q$ operations. However, a further economy can be made. We note that $e(-(k+q)/(2q)) = -e(-k/(2q))$. Thus

$$\begin{aligned}\widehat{f}(k) &= \widehat{g}(k) + e(-k/(2q))\widehat{h}(k), \\ \widehat{f}(k+q) &= \widehat{g}(k) - e(-k/(2q))\widehat{h}(k).\end{aligned}$$

The multiplication of $e(-k/(2q))$ by $\widehat{h}(k)$ is performed only once, and the result of this multiplication is either added to $\widehat{g}(k)$ or subtracted from it. Thus we determine both $\widehat{f}(k)$ and $\widehat{f}(k+q)$ in three arithmetic operations. This is done q times to obtain $\widehat{f}(k)$ for all k. Thus $N(2q) \le 2N(q) + 3q$. □

Corollary 2.11. *Let $N(q)$ be defined as in Theorem* 2.10*. Then $N(2^r) \le \frac{3}{2}r2^r$.*

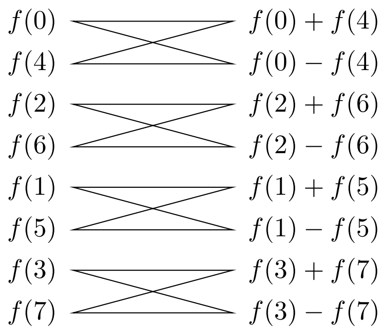
Proof. We note that $N(1) = 0$, since $\widehat{f}(k) = f(n)$ when f is an arithmetic function with period 1. Suppose that $N(2^r) \le \frac{3}{2}r2^r$. By Theorem 7 it follows that

$$N(2^{r+1}) \le 2N(2^r) + 3\cdot 2^r \le 3r2^r + 3\cdot 2^r = 3(r+1)2^r = \frac{3}{2}(r+1)2^{r+1}\,.$$

Thus the result follows by induction. □

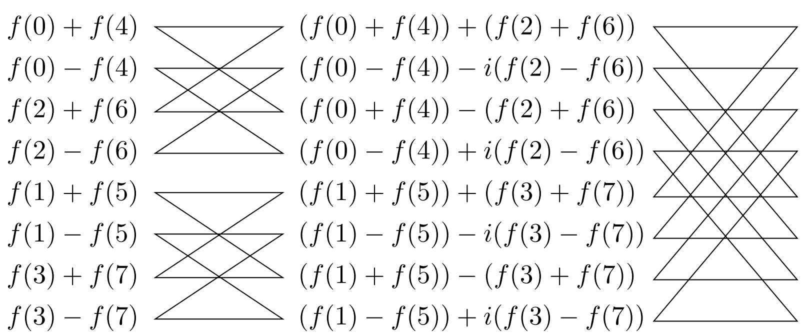
At this stage of our discussion, we do not have a precise algorithm, but it is fairly plausible that by making systematic use of binary expansions we should be able to compute all q Discrete Fourier Transforms in $O(q\log q)$ operations, when q is a power of 2. Let us consider how the calculation would proceed when $q = 8$. For brevity we let $\omega = e(-1/8)$. We start with a first column of eight numbers in the order shown below left, and form linear combinations to create a second column, shown below right.

(2.24)



Number the rows 0 through 7. We combine rows $2j$ and $2j+1$ first with a $+$ sign and then with a $-$ sign to form rows $2j$ and $2j+1$ in the second column. The sequence of mulipliers $1, -1, \dots, 1, -1$ applied to the second member in row j on the right can be thought of as $(\omega^4)^j$. Since we have no further use for the numbers in first column, the numbers in the second column can be stored in the same memory locations that had been used to store the first column. We next use the second column to form a third column

(2.25)



Here row $4j$ is combined with row $4j+2$ in two different ways to form new rows $4j$ and $4j+2$, and row $4j+1$ is combined with row $4j+3$ in two different ways to form new rows $4j+1$ and $4j+3$. Here the multiplier $1, -i, -1, i, \dots$ applied to the second term in row j is $(\omega^2)^j$. Finally, we use this third column to generate a fourth column

$$
\begin{aligned}
\widehat{f}(0) &= ((f(0)+f(4))+(f(2)+f(6)))+((f(1)+f(5))+(f(3)+f(7)))\\
\widehat{f}(1) &= ((f(0)-f(4))-i(f(2)-f(6)))+\omega((f(1)-f(5))-i(f(3)-f(7)))\\
\widehat{f}(2) &= ((f(0)+f(4))-(f(2)+f(6)))+\omega^2((f(1)+f(5))-(f(3)+f(7)))\\
\widehat{f}(3) &= (((f(0)-f(4))+i(f(2)-f(6)))+\omega^3((f(1)-f(5))+i(f(3)-f(7)))\\
\widehat{f}(4) &= ((f(0)+f(4))+(f(2)+f(6)))+\omega^4((f(1)+f(5))+(f(3)+f(7)))\\
\widehat{f}(5) &= ((f(0)-f(4))-i(f(2)-f(6)))+\omega^5((f(1)-f(5))-i(f(3)-f(7)))\\
\widehat{f}(6) &= ((f(0)+f(4))-(f(2)+f(6)))+\omega^6((f(1)+f(5))-(f(3)+f(7)))\\
\widehat{f}(7) &= ((f(0)-f(4))+i(f(2)-f(6)))+\omega^7((f(1)-f(5))+i(f(3)-f(7)))
\end{aligned}
$$

For $0 \le j < 4$ we combine row j with row $j+4$ in two different ways to form a new row j and row $j+4$. The multiplier attached to the second term in row j is ω^j. One may wonder at the peculiar order in which we listed the values of $f(n)$ to start this calculation. However, all becomes clear when one observes that the numbers 0 through 7 in binary are 000, 001, 010, 011, 100, 101, 110, 111. When the binary digits are written in reversed order, we have 000, 100, 010, 110, 001, 101, 011, 111. These are the binary expansions of $0, 4, 2, 6, 1, 5, 3, 7$, respectively. We could go on to construct an inductive proof that these observed patterns continue indefinitely. The calculation can be coded efficiently, so that the nominal saving of $O(q \log q)$ versus $O(q^2)$ is achieved without an overly severe penalty in the constants.

For many applications we can choose the q we want to work with, so we are free to take q to be a power of 2. For purposes of digitizing music, we want q to be of the order of 10^4 or 10^5, so $O(q \log q)$ is a huge savings on the naïve $O(q^2)$.

Suppose that f and g have period q and that we want to compute the convolution $(f * g)(n)$. This requires $\asymp q$ arithmetic operations for one particular n. If we repeat this for all q values of n, then we perform $\asymp q^2$ arithmetic operations. However, by the FFT we can compute $\widehat{f}(k)$ and $\widehat{g}(k)$ in $\asymp q \log q$ operations, the products $\widehat{f}(k)\widehat{g}(k)$ in q multiplications, and by a third application of FFT we recover $(f * g)(n)$ for all n in an additional $\asymp$ $q \log q$ operations. Thus FFT can be useful even when the Fourier transform is not the object of our interest. For example, when computing the coefficients of the product of two polynomials we are calculating a convolution. Thus the FFT can be used for rapid multiplication of polynomials of high degree.

Notes

2.2. The machinery of linear algebra, as discussed in Appendix L, enables an instructive insight concerning the Discrete Fourier Transform. Let $U = [u_{jk}]$ be a $q \times q$ matrix with entries

$$u_{jk} = \frac{e(-jk/q)}{\sqrt{q}}\,.$$

The inner product of the k^{th} column of this matrix with the ℓ^{th} column is

$$\sum_{j=1}^{q} \frac{e(-jk/q)}{\sqrt{q}} \cdot \frac{e(j\ell/q)}{\sqrt{q}} = \frac{1}{q}\sum_{j=1}^{q} e(j(\ell-k)/q) = \begin{cases} 1 & (k=\ell), \\ 0 & (\text{otherwise}). \end{cases}$$

That is, the columns of this matrix are orthonormal. Let $\boldsymbol{f} \in \mathbb{C}^q$ be a vector with coordinates $f(j)$. From the definition (2.6) of the Discrete Fourier Transform, we see that

$$U\boldsymbol{f} = \frac{1}{\sqrt{q}}\widehat{\boldsymbol{f}}\,.$$

But $\|U\boldsymbol{f}\| = \|\boldsymbol{f}\|$, so we obtain again Parseval's Identity (Corollary 2.8) for the Discrete Fourier transform. If we had put a factor $1/\sqrt{q}$ in our definition of the Discrete Fourier transform, then the act of taking the transform would be a unitary transformation, and all distances would be preserved (and there would be no factor q in Theorem 2.7). Some authors do this, and one can see the advantage of this. We instead adopted a convention by which the FFT is easier to discuss. As things stand, all distances are rescaled by the same factor.

In Appendix L.3 we use the DFT to diagonalize circulant matrices.

2.3. The Fast Fourier Transform was introduced by Cooley and Tukey (1965). Cooley and Tukey had both been involved with John von Neumann's project in the 1940s to build an early digital computer at the Institute for Advanced Study in Princeton. At that time, Tukey invented the term "bit", short for "binary digit". In the 1960s, for the purposes of a treaty with the Soviet Union, the U. S. needed to develop means to detect atomic explosions from a distance. Tukey, a statistician at Princeton University, became involved in this effort to analyze offshore seismic data. He had the critical insight to realize the computation could be greatly reduced, and joined forces with Cooley, a researcher at IBM, who was the first to program the algorithm and thus verified its feasibility. A secondary application at that time was for the acoustic detection of nuclear submarines. Other applications soon followed. The Cooley–Tukey paper revolutionized the theory of mathematical algorithms. It has been said that this paper has been cited more times than any other paper in the mathematical literature. The FFT is used routinely in optics, acoustics, quantum physics, telecommunications, signal processing, and image processing. Thus it is incorporated on a DVD, in JPEG and MP3 formats, and in the tomography of MRI and CAT scans. Our world would be very different indeed without this algorithm.

Over time, it has become recognized that the FFT was discovered, in whole or in part, many times. Gauss had a form of it in 1805, as did Runge in 1803, and it seems that Lagrange may have had it earlier still. It's a tangled (but interesting) tale. It's safe to say that Cooley and Tukey discovered it for the last time. For a more detailed treatment of the FFT, see Benedetto (1997, pp. 239–248).

Summability of Fourier Series

4.1. Cesàro summability of Fourier Series

When we define a number to be the value of a series, say

$$a = \sum_{n=0}^{\infty} a_n,$$

we think of the series as providing a formula for a. However, all that is being asserted is that the partial sums

$$s_N = \sum_{n=0}^{N} a_n \tag{4.1}$$

form a sequence of approximations to a. If the sequence of partial sums fails to have a limit, or if the limit is approached only very slowly, then sometimes something useful can be salvaged by forming averages of the partial sums,

$$\sigma_N = \frac{1}{N}\sum_{n=0}^{N-1} s_n = \frac{1}{N}\sum_{n=0}^{N-1}\sum_{k=0}^{n} a_k = \frac{1}{N}\sum_{k=0}^{N-1} a_k \sum_{n=k}^{N-1} 1 = \sum_{k=0}^{N-1}\Big(1 - \frac{k}{N}\Big) a_k\,. \tag{4.2}$$

We say that the series $\sum_{n=0}^{\infty} a_n$ is *Cesàro summable to* a, and write

$$\sum_{n=0}^{\infty} a_n = a \qquad (C)$$

if

$$\lim_{N\to\infty} \sigma_N = a\,.$$

Example 4.1. Consider the series

$$\sum_{n=0}^{\infty} (-1)^n\,.$$

The series does not converge, and indeed cannot converge, because its terms do not tend to 0. The sequence of its partial sums is $1, 0, 1, 0, \ldots$. Averages of these partial sums are $1, 1/2, 2/3, 1/2, 3/5, 1/2, 4/7, 1/2, \ldots$. In general, $\sigma_{2N} = 1/2$ and $\sigma_{2N+1} = (N+1)/(2N+1)$ and we see that $\sigma_N \to 1/2$ as $N \to \infty$. Accordingly, we write

$$\sum_{n=0}^{\infty}(-1)^n = \frac{1}{2} \qquad (C).$$

The partial sums $s_N(x)$ of a Fourier Series $\sum_{n=-\infty}^{\infty} \widehat{f}(n)e(nx)$ often do not converge very well, but when we form the corresponding Cesàro partial sum

$$\sigma_N(x) = \frac{s_0(x) + s_1(x) + \cdots + s_{N-1}(x)}{N} = \sum_{n=-N}^{N}(1 - |n|/N)\widehat{f}(n)e(nx) \tag{4.3}$$

the behavior is usually much better. By (3.11) and (3.12), $\sigma_N(x)$ can be expressed in trigonometric form,

$$\sigma_N(x) = \frac{1}{2}a_0 + \sum_{n=1}^{N}(1 - n/N)(a_n \cos 2\pi nx + b_n \sin 2\pi nx)\,. \tag{4.4}$$

Let

$$\Delta_N(x) = \sum_{n=-N}^{N}(1 - |n|/N)e(nx) = 1 + 2\sum_{n=1}^{N}(1 - n/N)\cos 2\pi nx\,. \tag{4.5}$$

This trigonometric polynomialis known as the *Fejér kernel.* In the same way that we used Lemma 3.14 to show that the partial sums $s_N(x)$ can be expressed as a convolution of f with the Dirichlet kernel D_N (recall (3.41)), we see by Lemma 3.14 that

$$\sigma_N(x) = (\Delta_N * f)(x) = \int_0^1 \Delta_N(u)f(x-u)\,du\,. \tag{4.6}$$

Our first order of business is to explore the properties of the Fejér kernel. In passing we note that the sums in (4.3) and (4.5) actually run from $-N+1$ to $N-1$, since the coefficient is 0 when $n = \pm N$. Thus $\Delta_N(x)$ is a trigonometric polynomial of degree $N-1$, and $\sigma_N(x)$ is a trigonometric polynomial of degree not exceeding $N-1$.

First we note that from (3.43) and (4.5) it is evident that

$$\int_0^1 \Delta_N(x)\,dx = 1 \tag{4.7}$$

for all N. Next we show that

$$\Delta_N(x) = \begin{cases} \dfrac{1}{N}\Big(\dfrac{\sin \pi Nx}{\sin \pi x}\Big)^2 & (x \notin \mathbb{Z}), \\ N & (x \in \mathbb{Z}). \end{cases} \tag{4.8}$$

First assume that x is not an integer. From the formula (0.2) for the sum of a geometric progression we see that

$$\sum_{n=0}^{N-1} e(nx) = \frac{1-e(Nx)}{1-e(x)} = \frac{e(Nx)-1}{e(x)-1}\,.$$

We factor out appropriate exponentials in the numerator and denominator to create expressions of the form $e(u)-e(-u)$. Thus the above is

$$= \frac{e(Nx/2)\big(e(Nx/2)-e(-Nx/2)\big)}{e(x/2)\big(e(x/2)-e(-x/2)\big)}\,.$$

On dividing the numerator and denominator by $2i$, we see that the above is

$$= e((N-1)x/2)\frac{\sin \pi N x}{\sin \pi x}\,. \tag{4.9}$$

Let z denote the common value of the left hand side and right hand side of this identity. Then $|z|^2$ is

$$\Big|\sum_{n=0}^{N-1} e(nx)\Big|^2 = \Big(\frac{\sin \pi N x}{\sin \pi x}\Big)^2\,.$$

We multiply out the left hand side to see that it is

$$\Big(\sum_{m=0}^{N-1} e(mx)\Big)\Big(\sum_{n=0}^{N-1} e(-nx)\Big) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} e((m-n)x)\,.$$

On collecting those m and n for which $m-n=k$, we see that the above is

$$= \sum_{k=-N+1}^{N-1} e(kx) \sum_{\substack{0\le m\le N-1\\ 0\le n\le N-1\\ m-n=k}} 1\,.$$

Here the inner sum is $N-|k|$, so on dividing both sides by N we obtain (4.8) when x is not an integer.

Since $\sin u \sim u$ as $u \to 0$, we deduce that

$$\lim_{x\to 0} \Delta_N(x) = \lim_{x\to 0} \frac{1}{N}\Big(\frac{\sin \pi N x}{\sin \pi x}\Big)^2 = N\,.$$

But $\Delta_N(x)$ is a trigonometric polynomial, and hence continuous, so $\Delta_N(0) = N$. Alternatively, from the definition (4.5) we see that

$$\begin{aligned}\Delta_N(0) &= 1 + 2\sum_{n=1}^{N-1}(1-n/N) = 1 + \frac{2}{N}\sum_{n=1}^{N-1} N-n \\ &= 1 + \frac{2}{N}\sum_{n=1}^{N-1} n = 1 + \frac{2}{N}\cdot\frac{N(N-1)}{2} = N\end{aligned}$$

by the formula (0.1) for the sum of an arithmetic progression. Thus (4.8) holds for all x.

The function $\sin N\pi x$ has simple zeros at $x = 0, 1/N, 2/N, \ldots$, and so $\Delta_N(x)$ has double zeros at these points, except for $x = 0$, since $\sin \pi x$ also vanishes there.

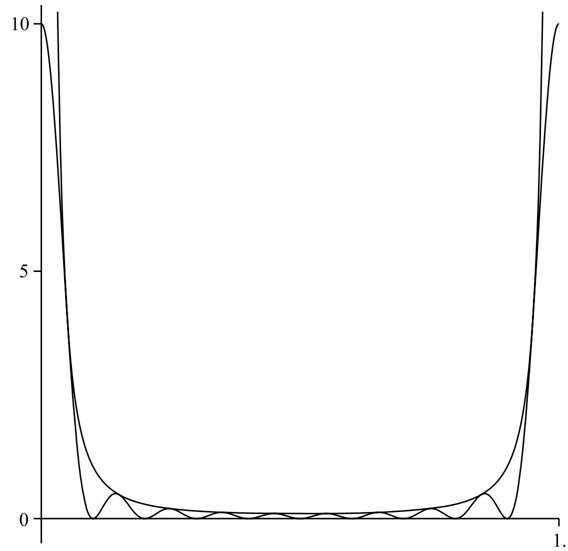


Figure 4.1. Graph of $\Delta_{10}(x)$ showing 9 double zeros and the envelope $1/(10\sin^2 \pi x)$.

Thus $\Delta_N(x)$ has $N-1$ double zeros, making a total of $2N-2$ zeros. Since $\Delta_N(x)$ is a trigonometric polynomial of degree $N-1$, this is the maximum number of zeros that it could have, in view of Theorem 3.15.

As $\Delta_N(x) \ge 0$ for all x and $\int_0^1 \Delta_N(x)\,dx = 1$, it follows that the convolution (4.6) is a weighted average of the values of f. Hence

$$|\sigma_N(x)| \le \max_x |f(x)|, \tag{4.10}$$

and if f is real-valued, then

$$\min_x f(x) \le \sigma_N(x) \le \max_x f(x) \tag{4.11}$$

for all x.

Since $\sin \pi x \ge 2x$ for $0 \le x \le 1/2$, we see that

$$\Delta_N(x) \le \min\Big(N, \frac{1}{4Nx^2}\Big) \tag{4.12}$$

for $-1/2 \le x \le 1/2$. Thus

$$\int_\delta^{1/2} \Delta_N(x)\,dx \le \int_\delta^{1/2} \frac{1}{4Nx^2}\,dx < \int_\delta^\infty \frac{1}{4Nx^2}\,dx = \frac{1}{4N\delta}\,. \tag{4.13}$$

If $\delta < 1/(2N)$, then we can do better by observing that

$$\int_\delta^{1/2} \Delta_N(x)\,dx \le \int_0^{1/2} \Delta_N(x)\,dx = 1/2.$$

Since $\Delta_N(x)$ is even, it follows that

$$\int_{1/2}^{1-\delta} \Delta_N(x)\,dx = \int_{\delta}^{1/2} \Delta_N(x)\,dx\,.$$

Hence

$$\int_{\delta}^{1-\delta} \Delta_N(x)\,dx \le \min\Big(1, \frac{1}{2N\delta}\Big) \tag{4.14}$$

for $0 < \delta \le 1/2$. We now put the properties of Δ_N to good use.

Theorem 4.1. *If f is a continuous function with period 1, then $\sigma_N(x)$ tends to $f(x)$ uniformly as $N \to \infty$.*

In other words, for any $\varepsilon > 0$, there is an N_0 such that if $N > N_0$, then

$$|f(x) - \sigma_N(x)| < \varepsilon \tag{4.15}$$

for all x.

Let $C(\mathbb{T})$ denote the set of continuous functions with period 1. For such functions we define a norm, called the *uniform norm*:

$$\|f\|_\infty = \max_{x\in\mathbb{T}} |f(x)|\,. \tag{4.16}$$

The distance between two members f, g of $C(\mathbb{T})$ is taken to be $\|f - g\|_\infty$. The import of Theorem 4.1 is that trigonometric polynomials are dense in $C(\mathbb{T})$: For any $f \in C(\mathbb{T})$ and any $\varepsilon > 0$, there is a trigonometric polynomial T such that $\|f - T\|_\infty < \varepsilon$.

Proof. From (4.6) and (4.7) we deduce that

$$f(x) - \sigma_N(x) = \int_0^1 \Delta_N(u)(f(x) - f(x-u))\,du\,.$$

We write this as

$$\int_{-\delta}^{\delta} + \int_{\delta}^{1-\delta} = I_1 + I_2\,.$$

Since f is continuous on $[0,1]$, it follows (see Theorem 0.5) that f is uniformly continuous. That is, for any $\varepsilon > 0$, there is a $\delta > 0$ such that for all x, $|f(x) - f(x-u)| < \varepsilon$ whenever $|u| \le \delta$. Hence if δ is chosen in this way, then

$$|I_1| \le \varepsilon \int_{-\delta}^{\delta} \Delta_N(u)\,du \le \varepsilon \int_0^1 \Delta_N(u)\,du = \varepsilon\,.$$

On the other hand, since f is continuous it follows that $|f(x)|$ is bounded, say $|f(x)| \le M$ for all x. Hence $|f(x) - f(x-u)| \le |f(x)| + |f(x-u)| \le 2M$, so

$$|I_2| \le 2M \int_{\delta}^{1-\delta} \Delta_N(u)\,du \le \frac{M}{N\delta}$$

by (4.14). Thus $I_2 \to 0$ as $N \to \infty$. In particular, $|I_2| < \varepsilon$ for all sufficiently large N. Thus

$$|f(x) - \sigma_N(x)| < 2\varepsilon$$

for all x and all sufficiently large N, which is the desired result. □

Corollary 4.2. *If $f \in C(\mathbb{T})$ and $\widehat{f}(n) = 0$ for all integers n, then $f(x) = 0$ for all x.*

Proof. We know that for any $|f(x) - \sigma_N(x)| < \varepsilon$ for all x, if N is sufficiently large. But if $\widehat{f}(n) = 0$ for all n, then $\sigma_N(x) = 0$ for all N and all x. Since $|f(x)| < \varepsilon$ for all x and ε is arbitrarily small, it follows that $f(x) = 0$ for all x. □

The following equivalent formulation is sometimes more convenient.

Corollary 4.3. *Suppose that $f \in C(\mathbb{T})$ and that $g \in C(\mathbb{T})$. If $\widehat{f}(n) = \widehat{g}(n)$ for all n, then $f(x) = g(x)$ for all x.*

Proof. Apply Corollary 4.2 to the function $f(x) - g(x)$, whose Fourier coefficients are $\widehat{f}(n) - \widehat{g}(n) = 0$ for all n. □

This has an immediate application to absolutely convergent trigonometric series. The following theorem sets the stage.

Theorem 4.4. *Suppose that*

$$\sum_{n=-\infty}^{\infty} |c_n| < \infty, \tag{4.17}$$

and put

$$g(x) = \sum_{n=-\infty}^{\infty} c_n e(nx)\,. \tag{4.18}$$

Then $g \in C(\mathbb{T})$, and $\widehat{g}(n) = c_n$ for all n.

Proof. The function g has period 1 because each term in its definition has period 1. Suppose that $\varepsilon > 0$ is given. By the convergence of the series (4.17) it follows that if N is a sufficiently large positive integer, then

$$\sum_{|n|>N} |c_n| < \varepsilon\,.$$

Let $s_N(x) = \sum_{n=-N}^{N} c_n e(nx)$. By the triangle inequality it follows that

$$|g(x) - s_N(x)| = \Big|\sum_{|n|>N} c_n e(nx)\Big| \le \sum_{|n|>N} |c_n| < \varepsilon \tag{4.19}$$

for all x. Since the partial sums $s_N(x)$ converge uniformly to $g(x)$, and since each term $c_n e(nx)$ is a continuous function of x, it follows by Theorem 0.5 that g is continuous.

Fix an integer m, and suppose (as we may) that $-N \le m \le N$. From (3.15) we see that

$$|\widehat{g}(m) - \widehat{s}_N(m)| \le \int_0^1 |g(x) - s_N(x)|\,dx < \int_0^1 \varepsilon\,dx = \varepsilon,$$

in view of (4.19). Now $s_N(x)$ is a trigonometric polynomial, and in (3.43) we established that the coefficients of a trigonometric polynomial are its Fourier coefficients.

Thus $\widehat{s}_N(m) = c_m$. Hence $|\widehat{g}(m) - c_m| < \varepsilon$. Since this is true for any $\varepsilon > 0$, it follows that $\widehat{g}(m) = c_m$, and the proof is complete. □

Example 4.2. Let $\delta(1), \delta(2), \ldots$ be a sequence of numbers tending monotonically to 0. Let $n_1, n_2, \ldots$ be a strictly increasing sequence of positive integers such that n_k is so large that $\delta(n_k) < 1/(2k^2)$. Put

$$f(x) = \sum_{k=1}^{\infty} \frac{\cos 2\pi n_k x}{k^2}\,.$$

By Theorem 4.4 we know that f is a continuous function, and that

$$\widehat{f}(n) = \begin{cases} \frac{1}{2k^2} & (n = \pm n_k \text{ for some } k), \\ 0 & (\text{otherwise}). \end{cases}$$

Thus there exist arbitrarily large n for which $\widehat{f}(n) > \delta(n)$. Here the $\delta(n)$ may tend to 0 slowly, in which case the n_k will increase very rapidly, and then $\widehat{f}(n) = 0$ for all but a very thin set of positive integers.

Suppose that $n_k = k^2$, so that

$$f(x) = \sum_{k=1}^{\infty} \frac{\cos 2\pi k^2 x}{k^2}\,. \tag{4.20}$$

This formula cannot be obtained from Theorem 3.17, since when the above is differentiated term-by-term we obtain a series whose coefficients do not tend to 0. Thus Theorem 4.4 expands our collection of convergent Fourier Series.

From Corollary 4.3 and Theorem 4.4 we obtain the following useful result.

Corollary 4.5. *Suppose that $f \in C(\mathbb{T})$. If $\sum_{n=-\infty}^{\infty} |\widehat{f}(n)| < \infty$, then*

$$f(x) = \sum_{n=-\infty}^{\infty} \widehat{f}(n) e(nx) \tag{4.21}$$

for all x.

Proof. Let $g(x)$ denote the sum on the right hand side above. By Theorem 4.4 we know that g is a continuous function with period 1, and that $\widehat{g}(n) = \widehat{f}(n)$ for all n. By Corollary 4.3 it follows that $f(x) = g(x)$ for all x. □

From Theorem 4.1 we know that any continuous function with period 1 can by uniformly approximated by a trigonometric polynomial. The corresponding result for algebraic polynomials is a famous result of of Weierstrass:

Corollary 4.6. (Weierstrass) *Suppose that f is continuous on the interval $[a, b]$. For any $\varepsilon > 0$ there is a polynomial $P(x) = \sum_{n=0}^{N} a_n x^n$ such that $|f(x) - P(x)| < \varepsilon$ uniformly for $x \in [a, b]$.*

Proof. Suppose first that $a = 0$ and $b = 1/2$. Put $f_1(x) = f(x)$ for $0 \le x \le 1/2$, and for $1/2 < x \le 1$ let $f_1(x)$ be linear with $f_1(1/2) = f(1/2)$ and $f_1(1) = f(0)$.

Then let $f_1(x)$ have period 1. By Theorem 4.1, there is a trigonometric polynomial $T(x) = \sum_{-N}^{N} t_n e(nx)$ such that $|f_1(x) - T(x)| < \varepsilon_1$ for all x. The power series

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

converges for all z, and uniformly so for $|z| \le R$, for any given R. Take $R = \pi N$. Then for any $\varepsilon_2 > 0$ there is a K such that

$$\Big| e^z - \sum_{k=0}^{K} \frac{z^k}{k!} \Big| < \varepsilon_2$$

uniformly for $|z| \le \pi N$. In particular, the above holds when $z = 2\pi i n x$, $-N \le n \le N$, and $0 \le x \le 1/2$. Thus

$$\Big| f(x) - \sum_{n=-N}^{N} t_n \sum_{k=0}^{K} \frac{(2\pi i n x)^k}{k!} \Big| < \varepsilon_1 + C\varepsilon_2$$

for $0 \le x \le 1/2$. Here $C = \sum_{-N}^{N} |t_n|$ may depend on ε_1, but we can take $\varepsilon_2 = \varepsilon_1/C$, and then the above is $< 2\varepsilon_1$. Thus we obtain a good uniform polynomial approximation to $f(x)$, uniformly for $0 \le x \le 1/2$.

Now consider a general interval $[a, b]$. We may assume that $a < b$. Let $\ell(x) = 2(b-a)x + a$. Thus $\ell(x)$ is a linear function with the property that $\ell(0) = a$ and $\ell(1/2) = b$. Put $f_1(x) = f(\ell(x))$ for $0 \le x \le 1/2$. Then f_1 is continuous, so there is a polynomial $P(x)$ such that $|f_1(x) - P(x)| < \varepsilon$ for $0 \le x \le 1/2$. On replacing x by $\ell^{-1}(x)$, we see that $|f_1(\ell^{-1}(x)) - P(\ell^{-1}(x))| < \varepsilon$ for $a \le x \le b$. But $f_1(\ell^{-1}(x)) = f(x)$, and $\ell^{-1}(x) = (x-a)/(2(b-a))$ is linear, so $P(\ell^{-1}(x))$ is a polynomial in x. Thus we have a good polynomial approximation to $f(x)$ for $a \le x \le b$. □

The proof just completed is a bit tortuous and unnatural. The Weierstrass uniform approximation theorem can be proved in many ways, and some of the proofs are very elegant. Our point is not that we have an attractive proof, but rather that Corollary 4.6 is an easy consequence of Theorem 4.1.

The same good properties of $\Delta_N(x)$ that gave rise to Theorem 4.1 can also be applied to an arbitrary function $f \in L^1(\mathbb{T})$.

Theorem 4.7. *Suppose that $f \in L^1(\mathbb{T})$. Then*

$$\lim_{N\to\infty} \int_0^1 |f(x) - \sigma_N(x)|\, dx = 0\,.$$

The relation above is referred to as *convergence in norm.* Since continuous functions are dense in $L^1(\mathbb{T})$ and since we can approximate a continuous function uniformly by a trigonometric polynomial, we already know that trigonometric polynomials are dense in $L^1(\mathbb{T})$. The point of the above is that $\sigma_N(x)$ gives a good approximation.

Proof. Since

$$\sigma_N(x) = \int_0^1 f(x-u)\Delta_N(u)\, du,$$

and since $\int_0^1 \Delta_N(u)\,du = 1$, it follows that

$$f(x) - \sigma_N(x) = \int_0^1 (f(x) - f(x-u))\Delta_N(u)\,du\,.$$

We take absolute values of both sides and integrate to see that

$$\int_0^1 |f(x) - \sigma_N(x)|\,dx = \int_0^1 \Big|\int_0^1 (f(x) - f(x-u))\Delta_N(u)\,du\Big|\,dx\,.$$

By the triangle inequality, the above is

$$\le \int_0^1 \int_0^1 |f(x) - f(x-u)|\Delta_N(u)\,du\,dx\,.$$

On interchanging the order of integration, we see that the above is

$$\begin{aligned} &= \int_0^1 \Delta_N(u)\Big(\int_0^1 |f(x) - f(x-u)|\,dx\Big)\,du \\ &= \int_{-\delta}^{\delta} + \int_{\delta}^{1-\delta} = I_1 + I_2, \end{aligned}$$

say. By Lemma 3.5 we know that for any $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\int_0^1 |f(x) - f(x-u)|\,dx < \varepsilon$$

whenever $|u| < \delta$. With δ chosen in this way, we find that

$$|I_1| \le \varepsilon \int_{-\delta}^{\delta} \Delta_N(u)\,du \le \varepsilon \int_0^1 \Delta_N(u)\,du = \varepsilon\,.$$

As for I_2, we observe that

$$\int_0^1 |f(x) - f(x-u)|\,dx \le \int_0^1 |f(x)|\,dx + \int_0^1 |f(x-u)|\,dx = 2\|f\|_1\,.$$

Hence by (4.14),

$$|I_2| \le 2\|f\|_1 \int_{\delta}^{1-\delta} \Delta_N(u)\,du \le \frac{\|f\|_1}{N\delta}\,.$$

Hence $I_2 \to 0$ as $N \to \infty$. On combining our estimates, we find that

$$\int_0^1 |f(x) - \sigma_N(x)|\,dx < 2\varepsilon$$

for all sufficiently large N. Thus we have the stated result. □

In (3.5) we showed that the functions $e(nx)$ are orthonormal. We now show that this orthonormal system is complete in the sense that if f is orthogonal to all the functions $e(nx)$, then f is 0 almost everywhere.

Corollary 4.8. *If $f \in L^1(\mathbb{T})$, then $\int_0^1 |f(x)|\,dx = 0$ if and only if $\widehat{f}(n) = 0$ for all integers n.*

Proof. If $\int_0^1 |f(x)|\,dx = 0$, then $\widehat{f}(n) = 0$ for all n by inequality (3.15). Suppose conversely that $\widehat{f}(n) = 0$ for all n. Then $\sigma_N(x) = 0$ for all N and all x. From Theorem 4.7 we know that

$$\int_0^1 |f(x) - \sigma_N(x)|\,dx$$

tends to 0 as N tends to infinity. Thus $\int_0^1 |f(x)|\,dx$ tends to 0 as N tends to infinity. Since this quantity is independent of N, in order to tend to 0 it must be 0. ☐

Suppose that f is a measurable function with period 1. For each integer n, let

$$\mathcal{S}_n = \{x \in [0,1) \,:\, 2^{n-1} < |f(x)| \le 2^n\}\,. \tag{4.22}$$

We recall that the *measure* of a set is its length. For example, if a set $\mathcal{A}$ is a union of some disjoint intervals, then

$$\operatorname{meas} \mathcal{A} = \int_{\mathcal{A}} 1\,dx \tag{4.23}$$

is the sum of the lengths of those intervals. The structure of $\mathcal{A}$ may be more complicated than this, but this is an issue of measure theory that need not concern us, and (4.23) is true for any measurable set. Thus

$$\int_0^1 |f(x)|\,dx = \sum_{n=-\infty}^{\infty} \int_{\mathcal{S}_n} |f(x)|\,dx \le \sum_{n=-\infty}^{\infty} \int_{\mathcal{S}_n} 2^n\,dx = \sum_{n=-\infty}^{\infty} 2^n \operatorname{meas} \mathcal{S}_n = E,$$

say. Similarly,

$$\int_0^1 |f(x)|\,dx = \sum_{n=-\infty}^{\infty} \int_{\mathcal{S}_n} |f(x)|\,dx \ge \sum_{n=-\infty}^{\infty} \int_{\mathcal{S}_n} 2^{n-1}\,dx = \sum_{n=-\infty}^{\infty} 2^{n-1} \operatorname{meas} \mathcal{S}_n = E/2.$$

Thus $E \asymp \|f\|_1$, and in particular $E < \infty$ if and only if $\|f\|_1 < \infty$. Let

$$\mathcal{S} = \bigcup_{n=-\infty}^{\infty} \mathcal{S}_n, \qquad \mathcal{S}^* = \{x \in [0,1) \,:\, f(x) = 0\}\,. \tag{4.24}$$

Here $\mathcal{S}$ is the set of x for which $f(x) \ne 0$, and hence $\mathcal{S} \cup \mathcal{S}^* = [0,1)$. Consequently,

$$1 = \operatorname{meas} \mathcal{S}^* + \operatorname{meas} \mathcal{S} = \operatorname{meas} \mathcal{S}^* + \sum_{n=-\infty}^{\infty} \operatorname{meas} \mathcal{S}_n\,. \tag{4.25}$$

If $\|f\|_1 = 0$, then $E = 0$, so $\operatorname{meas} \mathcal{S}_n = 0$ for all n, so $\operatorname{meas} \mathcal{S} = 0$. Thus we say that $f(x) = 0$ for almost all x, which simply means that the set of exceptions has Lebesgue measure 0. Conversely, if $\operatorname{meas} \mathcal{S}^* = 1$, then $\operatorname{meas} \mathcal{S} = 0$, so $\operatorname{meas} \mathcal{S}_n = 0$ for all n, so $E = 0$, so $\|f\|_1 = 0$. Thus Corollary 4.8 can be expressed as follows.

Corollary 4.9. *Suppose that $f \in L^1(\mathbb{T})$. Then the following three assertions are equivalent:*

(a) $\displaystyle\int_0^1 |f(x)|\,dx = 0$;

(b) $\widehat{f}(n) = 0$ *for all integers* n;

(c) $f(x) = 0$ *for almost all real* x.

Another way to look at the above is to say that distinct functions have distinct sets of Fourier coefficients. That is, by applying the above to $f(x) - g(x)$ we obtain

Corollary 4.10. *Suppose that f and g are functions in $L^1(\mathbb{T})$. Then the following three assertions are equivalent:*

(a) $\displaystyle\int_0^1 |f(x) - g(x)|\, dx = 0$;

(b) $\widehat{f}(n) = \widehat{g}(n)$ *for all integers n;*

(c) $f(x) = g(x)$ *for almost all real x.*

In the situation of Corollary 4.9, we cannot deduce that $f(x) = 0$ for all x, but we can deduce that $f(x) = 0$ at points of continuity, in view of the following simple observation.

Theorem 4.11. *If $f \in L^1(\mathbb{T})$, if f is continuous at x_0, and if $f(x_0) \neq 0$, then*

$$\int_{x_0-\delta}^{x_0+\delta} |f(x)|\, dx > 0 \tag{4.26}$$

for all $\delta > 0$.

Proof. Let $a = f(x_0)$. Since f is continuous at x_0, for any $\varepsilon > 0$ there is a $\delta > 0$ such that $|f(x) - a| < \varepsilon$ for $x_0 - \delta < x < x_0 + \delta$. Take $\varepsilon = |a|/2$. Then by the triangle inequality, $|f(x)| = |f(x) - a + a| \geq |a| - |f(x) - a| \geq |a|/2$ for $x_0 - \delta < x < x_0 + \delta$. Hence the integral in (4.26) is $\geq |a|\delta$ for all sufficiently small δ. Thus the integral is positive for all $\delta > 0$. □

In particular, if $f \in C(\mathbb{T})$, then $\widehat{f}(n) = 0$ for all n if and only if $f(x) = 0$ for all x. As a second application of the above principle, we derive a useful extension of Corollary 4.5.

Corollary 4.12. *Suppose that $f \in L^1(\mathbb{T})$, and that $\sum_{n=-\infty}^{\infty} |\widehat{f}(n)| < \infty$. Then*

$$f(x) = \sum_{-\infty}^{\infty} \widehat{f}(n) e(nx) \tag{4.27}$$

for any x at which f is continuous.

Proof. We let $g(x)$ be the sum on the right hand side above. By Theorem 4.4 we know that g is a continuous function with period 1, and that $\widehat{g}(n) = \widehat{f}(n)$ for all n. By Corollary 4.10 it follows that $\int_0^1 |f(x) - g(x)|\, dx = 0$. Hence by Theorem 4.11, $f(x) = g(x)$ at any point where f is continuous. □

Theorem 4.7 is fundamental, as it allows one to show—in some situations—that if something is true for trigonometric polynomials, then it is also true for general functions. The following is an example of this line of reasoning.

Theorem 4.13. (Fejér's Lemma) *Suppose that $f \in L^1(\mathbb{T})$, that g is a bounded function with period 1. Then*

$$\lim_{m\to\infty} \int_0^1 f(x) g(mx)\, dx = \widehat{f}(0)\widehat{g}(0)\,.$$

We note that the Riemann–Lebesgue Lemma (Theorem 3.6) is the special case $g(x) = e(-x)$ of the above.

Proof. Suppose that N is so large that $\int_0^1 |f(x) - \sigma_N(x)|\,dx < \varepsilon$, and let C be a constant chosen so that $|g(x)| \le C$ for all x. Then by the triangle inequality

$$\left| \int_0^1 f(x)g(mx)\,dx - \widehat{f}(0)\widehat{g}(0) \right| \tag{4.28}$$

$$= \left| \int_0^1 \sigma_N(x)g(mx)\,dx - \widehat{f}(0)\widehat{g}(0) + \int_0^1 (f(x) - \sigma_N(x))g(mx)\,dx \right| \tag{4.29}$$

$$\le \left| \int_0^1 \sigma_N(x)g(mx)\,dx - \widehat{f}(0)\widehat{g}(0) \right| + \left| \int_0^1 (f(x) - \sigma_N(x))g(mx)\,dx \right| .$$

Again by the triangle inequality, the last term is

$$\le \int_0^1 |f(x) - \sigma_N(x)||g(mx)|\,dx \le C \int_0^1 |f(x) - \sigma_N(x)|\,dx < C\varepsilon\,. \tag{4.30}$$

Now

$$\int_0^1 \sigma_N(x)g(mx)\,dx = \sum_{n=-N}^{N} (1 - |n|/N)\widehat{f}(n) \int_0^1 e(nx)g(mx)\,dx,$$

which by Theorem 3.2 (d) is

$$= \sum_{\substack{-N<n<N \\ m|n}} (1 - |n|/N)\widehat{f}(n)\widehat{g}(-n/m)\,.$$

If $m \ge N$, then the only multiple of m in the interval $(-N, N)$ is $n = 0$. Hence the above is exactly $\widehat{f}(0)\widehat{g}(0)$ for all $m \ge N$. The stated result now follows by combining this with (4.28) and (4.30). □

Before continuing, we make a comment about notation: We let $f(a^-)$ denote the limit of $f(x)$ as x approaches a from below, and $f(a^+)$ denote the limit of $f(x)$ as x approaches a from above:

$$f(a^-) = \lim_{x\to a^-} f(x), \qquad f(a^+) = \lim_{x\to a^+} f(x)\,. \tag{4.31}$$

Theorem 4.14. (Fejér) *Suppose that $f \in L^1(\mathbb{T})$. If $f(x^-)$ and $f(x^+)$ both exist and are finite, then*

$$\lim_{N\to\infty} \sigma_N(x) = \frac{f(x^-) + f(x^+)}{2}\,.$$

Note that the value of $f(x)$ is irrelevant. However, to say that f is continuous at x is equivalent to saying that $f(x^-) = f(x^+) = f(x)$, and in that case the above asserts that $\sigma_N(x)$ tends to $f(x)$.

Proof. We write

$$\sigma_N(x) = \int_0^1 f(x-u)\Delta_N(u)\,du = \int_0^\delta + \int_\delta^{1-\delta} + \int_{1-\delta}^1 = I_1 + I_2 + I_3,$$

say. Here we suppose that δ has been chosen small enough to ensure that $|f(x-u)-f(x^-)|<\varepsilon$ uniformly for $0\le u\le\delta$ and $|f(x-u)-f(x^+)|<\varepsilon$ uniformly for $1-\delta\le u\le 1$. We write

$$I_1=\int_0^\delta (f(x-u)-f(x^-))\Delta_N(u)\,du+\ f(x^-)\int_0^\delta \Delta_N(u)\,du=T_1+T_1',$$

say. Now

$$|T_1|\le\int_0^\delta |f(x-u)-f(x^-)|\Delta_N(u)\,du<\varepsilon\int_0^\delta \Delta_N(u)\,du\le\varepsilon\int_{-1/2}^{1/2}\Delta_N(u)\,du=\varepsilon$$

by (4.7). From (4.7) and (4.13) we deduce that

$$\frac{1}{2}-\frac{1}{4N\delta}\le\int_0^\delta \Delta_N(u)\,du\le\frac{1}{2}\,.$$

Hence

$$\Big|T_1'-\frac{f(x^-)}{2}\Big|\le\frac{|f(x^-)|}{4N\delta}\,.$$

By (4.12) we see that

$$|I_2|\le\frac{1}{4N\delta^2}\int_\delta^{1-\delta}|f(x-u)|\,du\le\frac{1}{4N\delta^2}\int_0^1|f(u)|\,du\,.$$

We treat I_3 in the same way as I_1, with terms T_3 and T_3'. Thus

$$\sigma_N(x)=T_1+T_1'+I_2+T_3+T_3'\,.$$

Here $|T_1|<\varepsilon$ and $|T_3|<\varepsilon$ for all N, while $T_1'\to f(x^-)/2$, $I_2\to 0$, and $T_3'\to f(x^+)/2$ as $N\to\infty$. Hence $|\sigma_N(x)-(f(x^-)+f(x^+)/2|<3\varepsilon$ for all sufficiently large N so the proof is complete. □

As was discussed in §0.3, if f is a Lebesgue-integrable function, then x is said to be a *Lebesgue point* of f if

$$\lim_{h\to0^+}\frac{1}{h}\int_{x-h}^{x+h}|f(u)-f(x)|\,du=0\,.$$

Theorem 4.15. *If $f\in L^1(\mathbb{T})$, and if x is a Lebesgue point of f, then $\sigma_N(x)\to f(x)$ as $N\to\infty$.*

Proof. We know that

$$f(x)-\sigma_N(x)=\int_0^1 (f(x)-f(x-u))\Delta_N(u)\,du\,.$$

Let $R_N(x) = \min(N, 1/(N\|x\|^2))$, so that $\Delta_N(x) \le R_N(x)$ for all x by (4.12). Thus by the triangle inequality,

$$
\begin{aligned}
|f(x) - \sigma_N(x)| &\le \int_0^1 |f(x) - f(x-u)| R_N(u)\, du \\
&= \int_0^{1/2} \big(|f(x) - f(x-u)| + |f(x) - f(x+u)|\big) R_N(u)\, du \\
&= N \int_0^{1/N} \big(|f(x) - f(x-u)| + |f(x) - f(x+u)|\big)\, du \qquad (4.32)\\
&\quad + \frac{1}{N} \int_{1/N}^{1/2} \big(|f(x) - f(x-u)| + |f(x) - f(x+u)|\big) \frac{du}{u^2}\, . \qquad (4.33)
\end{aligned}
$$

Let

$$I(h) = \int_0^h |f(x) - f(x-u)| + |f(x) - f(x+u)|\, du\, .$$

By integrating by parts we see that the term (4.33) is

$$= \frac{4}{N} I(1/2) - N I(1/N) + \frac{2}{N} \int_{1/N}^{1/2} I(u)\, \frac{du}{u^3}\, .$$

The term (4.32) is $NI(1/N)$, so on adding these two terms we deduce that

$$|f(x) - \sigma_N(x)| \le \frac{4}{N} I(1/2) + \frac{2}{N} \int_{1/N}^{1/2} I(u)\, \frac{du}{u^3}\, . \qquad (4.34)$$

Suppose that x is a Lebesgue point of f. Then for any $\varepsilon > 0$ there is a $\delta > 0$ such that $I(h) < \varepsilon h$ when $0 \le h \le \delta$. Clearly $I(1/2) \le 2|f(x)| + \|f\|_1$. Thus the first quantity on the right above is small if N is large. We write the remaining expression above as

$$\frac{2}{N} \int_{1/N}^{\delta} I(u)\, \frac{du}{u^3} + \frac{2}{N} \int_{\delta}^{1/2} I(u)\, \frac{du}{u^3} = T_1 + T_2,$$

say. Since $I(u) < \varepsilon u$ for $u < \delta$, we see that

$$T_1 < \frac{2\varepsilon}{N} \int_{1/N}^{\delta} u^{-2}\, du < \frac{2\varepsilon}{N} \int_{1/N}^{\infty} u^{-2}\, du = 2\varepsilon\, .$$

Since $I(u) \le I(1/2)$ for $u \le 1/2$, we see that

$$T_2 \le \frac{2I(1/2)}{N} \int_{\delta}^{1/2} u^{-3}\, du < \frac{2I(1/2)}{N} \int_{\delta}^{\infty} u^{-3}\, du = \frac{I(1/2)}{N\delta^2},$$

which is small if N is large. □

Corollary 4.16. *If $f \in L^1(\mathbb{T})$, then for almost all x, $\sigma_N(x) \to f(x)$ as $N \to \infty$.*

Proof. As was noted in §0.3, almost all x are Lebesgue points of f. □

Exercises

1. Suppose that $f \in L^1(\mathbb{T})$.
   (a) Show that
$$\int_0^1 (1 \pm \cos 2\pi nx) f(x)\, dx = \frac{1}{2}a_0 \pm \frac{1}{2}a_n$$
   in the usual trigonometric notation.
   (b) Assuming that $f(x) \geq 0$ for all x, explain why the integrand above is everywhere nonnegative.
   (c) Deduce that if $f(x) \geq 0$ for all x, then $|a_n| \leq a_0$ for all positive n.
   (d) Show that a_n can be nearly as big as a_0 by considering $f(x) = \Delta_N(x)$ with N large compared with n.
2. Suppose that f is a bounded function with period 1 and that $g \in L^1(\mathbb{T})$. Explain how you know that
$$\sum_{n=-N}^{N} \Big(1 - \frac{|n|}{N}\Big) \widehat{f}(n)\widehat{g}(n)e(n\theta) \longrightarrow (f * g)(\theta)$$
   uniformly in θ. (Hint: Recall Theorem 3.13.)
3. Suppose that f is defined on the real line, and put
$$f_e(x) = \frac{f(x) + f(-x)}{2}, \qquad f_o(x) = \frac{f(x) - f(-x)}{2}.$$
   (a) Show that f_e is an even function and that f_o is an odd function.
   (b) Show that $f(x) = f_e(x) + f_o(x)$ for all x.
   (c) Suppose that $f(x) = g(x) + h(x)$ for all x, that g is an even function, and that h is an odd function. Show that $g(x) = f_e(x)$ and that $h(x) = f_o(x)$ for all x. We call f_e the *even part* of f, and f_o the *odd part* of f.
   (d) Show that if $f \in L^1(\mathbb{T})$, then
$$\widehat{f_e}(n) = \frac{\widehat{f}(n) + \widehat{f}(-n)}{2}, \qquad \widehat{f_o}(n) = \frac{\widehat{f}(n) - \widehat{f}(-n)}{2}.$$
   (e) Suppose that $f \in L^1(\mathbb{T})$. Show that if $\widehat{f}(n)$ is an even sequence, then $\widehat{f_o}(n) = 0$ for all n. Explain why it follows that there is an even function $f_e \in L^1(\mathbb{T})$ such that $f(x) = f_e(x)$ for almost all x.
   (f) Suppose that $f \in L^1(\mathbb{T})$. Show that if $\widehat{f}(n)$ is an odd sequence, then $\widehat{f_e}(n) = 0$ for all n. Explain why it follows that there is an odd function $f_o \in L^1(\mathbb{T})$ such that $f(x) = f_o(x)$ for almost all x. Note that these last two results give a partial converse of Corollary 3.3.
4. (a) For real x let $f(x) = 1/(x-1)$ when $x \neq 1$, and put $f(1) = 0$. Express f as the sum of an even function and an odd function.
   (b) Let f have period 1, $f(x) = x^2$ for $0 < x < 1$, and $f(0) = 1/2$. Express f as the sum of an even function and an odd function.

5. Suppose that $f \in L^1(\mathbb{T})$, and put
$$f_r(x) = \frac{f(x) + \overline{f(x)}}{2} = \operatorname{Re} f(x), \qquad f_i(x) = \frac{f(x) - \overline{f(x)}}{2i} = \operatorname{Im} f(x)\,.$$
   (a) Show that $f(x) = f_r(x) + if_i(x)$ for all x.
   (b) Show that
$$\widehat{f_r}(n) = \frac{\widehat{f}(n) + \overline{\widehat{f}(-n)}}{2}, \qquad \widehat{f_i}(n) = \frac{\widehat{f}(n) - \overline{\widehat{f}(-n)}}{2i}\,.$$
   (c) Conclude that if $\widehat{f}(-n) = \overline{\widehat{f}(n)}$ for all n, then f is almost a real-valued function in the sense that there is a real-valued function, namely f_r, such that $f(x) = f_r(x)$ for almost all x. Note that this gives a partial converse of Corollary 3.4.
6. Suppose that $f \in L^1(\mathbb{T})$. Show that if $\widehat{f}(n) = 0$ for all but finitely many n, then there is a trigonometric polynomial T such that $f(x) = T(x)$ for almost all x.
7. (a) Show that if $f \in C(\mathbb{T})$ and $\widehat{f}(-n) = \widehat{f}(n)$ for all n, then f is an even function.
   (b) Show that if $f \in C(\mathbb{T})$ and $\widehat{f}(-n) = -\widehat{f}(n)$ for all n, then f is an odd function.
   (c) Show that if $f \in C(\mathbb{T})$ and $\widehat{f}(-n) = \overline{\widehat{f}(n)}$ for all n, then f is a real-valued function.
8. Suppose that $f \in L^1(\mathbb{T})$, that n is an integer for which $|n|$ is large, set $N = |n| - 1$, and put $g(x) = f(x) - \sigma_N(f;x)$.
   (a) Explain why $\|g\|_1$ is small.
   (b) Explain why $|\widehat{g}(n)|$ is small.
   (c) Show that $\widehat{f}(n) = \widehat{g}(n)$.
   (d) Thus give a second proof of the Riemann–Lebesgue Lemma (Theorem 3.6). Of course, both proofs depend on Lemma 3.5.
9. Suppose that $p > 0$ and that $M_p = \int_a^b |f(x)|^p\,dx < \infty$.
   (a) Show that for $V > 0$,
$$\operatorname{meas}\{x \in [a,b] \,:\, |f(x)| \ge V\} \le \frac{M_p}{V^p}\,.$$
   (b) Show that $\operatorname{meas}\{x \in [a,b] \,:\, |f(x)| \ge V\} = o(1/V^p)$ as $V \to \infty$. (For a discussion of the "little oh" notation, see Appendix O.)

   If there is a constant C such that
$$\operatorname{meas}\{x \in [a,b] \,:\, |f(x)| \ge V\} \le C/V^p \tag{4.35}$$
   for all $V > 0$, then we say that f is of *weak type* L^p.
   (c) Suppose that $a < b$ and $p > 0$ are given. Give an example of a function f such that f is of weak type L^p on $[a,b]$, but $\int_a^b |f(x)|^p\,dx = \infty$.
   (d) Suppose that $a < b$, that $0 < p_2 < p_1$, and that f is of weak type L^{p_1} on $[a,b]$. Show that $\int_a^b |f(x)|^{p_2}\,dx < \infty$.

10. Suppose that $f \in L^1(\mathbb{T})$. Show that if f is continuous at x, then

$$|f(x)| \le \sum_{n=-\infty}^{\infty} |\widehat{f}(n)|.$$

(The sum on the right may diverge, but this is analogous to (2.15) for the Discrete Fourier Transform.)

11. (a) Show that $e(\alpha) - e(\beta) = 2ie((\alpha+\beta)/2)\sin \pi(\alpha-\beta)$.
    (b) Show that $e(\alpha) + e(\beta) = 2e((\alpha+\beta)/2)\cos \pi(\alpha-\beta)$.

12. Show that

$$\frac{1}{1-e(x)} = \frac{1}{2} + \frac{i}{2}\cot \pi x\,.$$

13. We define the *conjugate Fejér kernel* (plotted in Figure 4.2) to be

$$\widetilde{\Delta}_N(x) = \sum_{n=1}^{N-1}\Big(1-\frac{n}{N}\Big)\big(-ie(nx)+ie(-nx)\big) = 2\sum_{n=1}^{N-1}\Big(1-\frac{n}{N}\Big)\sin 2\pi nx\,. \tag{4.36}$$

(a) Let $s_N(x) = 1 + e(x) + e(2x) + \cdots + e((N-1)x)$. Show that

$$s_N(x) = \frac{1-e(Nx)}{1-e(x)}\,.$$

(b) Put

$$\sigma_N(x) = \sum_{n=0}^{N-1}(N-n)e(nx)\,.$$

Note that $\sigma_N = s_1(x) + \cdots + s_N(x)$. Show that

$$\sigma_N(x) = \frac{N}{1-e(x)} - e(x)\frac{1-e(Nx)}{(1-e(x))^2}\,.$$

(c) Show that the above is

$$= N\Big(\frac{1}{2}+\frac{i}{2}\cot \pi x\Big) - \frac{i}{2}e(Nx/2)\frac{\sin \pi x}{(\sin \pi x)^2}\,.$$

(d) Take imaginary parts and multiply by $2/N$ to deduce that

$$\widetilde{\Delta}_N(x) = \cot \pi x - \frac{\cos \pi Nx\, \sin \pi Nx}{N(\sin \pi x)^2}\,.$$

(e) Conclude that

$$\widetilde{\Delta}_N(x) = \cot \pi x - \frac{\sin 2\pi Nx}{2N(\sin \pi x)^2}\,. \tag{4.37}$$

14. (a) Show that if n is an integer, $0 < n < 2N$, then $\widetilde{\Delta}_N(n/(2N)) = \cot \pi n/(2N)$.
    (b) Show that if n is an integer, $0 < n < 2N$, then

$$\widetilde{\Delta}'_N(n/(2N)) = \begin{cases} 0 & (n \text{ odd}), \\ -2\pi \csc^2 \pi n/(2N) & (n \text{ even}). \end{cases}$$

(c) Show that if $0 < x < 1$, then

$$\lim_{N\to\infty} \widetilde{\Delta}_N(x) = \cot \pi x\,.$$

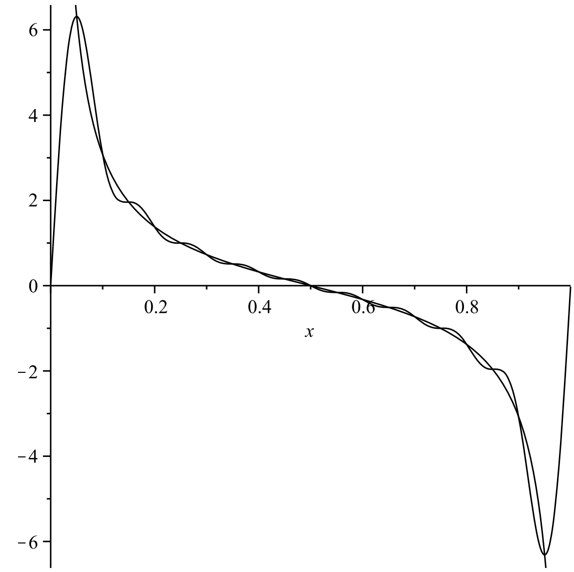


Figure 4.2. The conjugate Fejér kernel $\widetilde{\Delta}_{10}(x)$ and $\cot \pi x$.

(d) Show that for any integer n,

$$\lim_{N\to\infty} \widehat{\widetilde{\Delta}}_N(n) = \begin{cases} -i & (n>0), \\ 0 & (n=0), \\ i & (n<0). \end{cases}$$

Let $K(x) = \cot \pi x$. In some sense,

$$\widehat{K}(n) = \begin{cases} -i & (n>0), \\ 0 & (n=0), \\ i & (n<0). \end{cases} \tag{4.38}$$

Since these Fourier coefficients do not tend to 0 it is clear at once (if it were not already clear) that $K \notin L^1(\mathbb{T})$. Nevertheless, one might hope that if $f \in L^1(\mathbb{T})$, then

$$\int_0^1 K(u) f(x-u)\, du = \widetilde{f}(x), \tag{4.39}$$

the conjugate function of f, whose Fourier coefficients are

$$\widehat{\widetilde{f}}(n) = \begin{cases} -i\widehat{f}(n) & (n>0), \\ 0 & (n=0), \\ i\widehat{f}(n) & (n<0). \end{cases}$$

There are problems here, since the integral (4.39) is unlikely to converge, although

$$\lim_{\varepsilon\to 0^+} \int_\varepsilon^{1-\varepsilon} K(u) f(x-u)\, du$$

will exist if f is smooth enough at x. Still, it is known that there exist $f \in L^1(\mathbb{T})$ such that $\widehat{f} \notin L^1(\mathbb{T})$. In the next several exercises, we show that (4.38) is in some sense true.

15. Explain why
$$\int_{\varepsilon}^{1-\varepsilon} (\cot \pi u) \cos(2\pi nu\, du = 0$$
for all integers n and all $\varepsilon > 0$.

16. Put
$$g_1(x) = \frac{1}{\pi} \sum_{k=-K}^{K} \frac{1}{x-k},$$
$$g_2(x) = \frac{1}{\pi} \sum_{k>K} \Big(\frac{1}{x-k} + \frac{1}{x+k} \Big).$$
Then by (T.88) we know that $g_1(x) + g_2(x) = \cot \pi x$.
   (a) Show that
$$\int_{-1/2}^{1/2} g_1(u) \sin 2\pi nu\, du = \frac{1}{\pi} \int_{-K-1/2}^{K+1/2} \frac{2\pi nu}{u}\, du\,.$$
   (b) Suppose that $n > 0$. Show that
$$\int_{-K-1/2}^{K+1/2} \frac{\sin 2\pi nu}{u}\, du = \int_{-\pi n(2K+1)}^{\pi n(2K+1)} \frac{\sin v}{v}\, dv\,.$$
   (c) Show that if $0 < a < b$, then
$$\int_a^b \frac{\sin v}{v}\, dv = \frac{\cos a}{a} - \frac{\cos b}{b} - \int_a^b \frac{\cos v}{v^2}\, dv\,.$$
   (d) Deduce that if $1 \le a < b$, then
$$\int_a^b \frac{\sin v}{v}\, dv = O(1/a)\,.$$
   (e) Deduce that
$$\lim_{V\to\infty} \int_{-V}^{V} \frac{\sin v}{v}\, dv$$
exists.
   (f) Show that
$$\frac{1}{x-k} + \frac{1}{x+k} = O\big(1/k^2\big)$$
if $|k| \ge 1$ and $-1/2 \le x \le 1/2$.
   (g) Deduce that $g_2(x) = O(1/K)$ for $-1/2 \le x \le 1/2$.
   (h) Deduce that $\int_{-1/2}^{1/2} g_2(u) \sin 2\pi nu\, du = O(1/K)$.
   (i) Conclude that if n is a positive integer, then
$$\int_0^1 (\cot \pi u) \sin 2\pi nu\, du = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin v}{v}\, dv\,.$$

17. Let $f(x) = \csc x - 1/(\pi x)$. Note that by (T.89) this is a bounded function in the interval $-1/2 \le x \le 1/2$. Let

$$I = \int_{-1/2}^{1/2} f(x) \sin(2N+1)\pi x \, dx .$$

(a) Show that

$$I = \int_{-1/2}^{1/2} f'(x) \frac{\cos(2N+1)\pi x}{(2N+1)\pi} \, dx .$$

(b) By term-by-term differentiation of (T.89) show that $f'(x)$ is bounded for $-1/2 \le x \le 1/2$. Deduce that $I = O(1/N)$.

(c) Show that

$$I = \int_{-1/2}^{1/2} D_N(x) \, dx - \frac{1}{pi} \int_{-1/2}^{1/2} \frac{\sin(2N+1)\pi x}{x} \, dx .$$

(d) Show that

$$\int_{-1/2}^{1/2} \frac{\sin(2N+1)\pi x}{x} \, dx = \int_{(-N+1/2)\pi}^{(N+1/2)\pi} \frac{\sin v}{v} \, dv .$$

(e) Deduce that

$$\int_{-\infty}^{\infty} \frac{\sin v}{v} \, dv = \pi .$$

(f) Deduce that

$$\int_0^1 (\cot \pi x) \sin 2\pi n x \, dx = 1$$

for all positive integers n.

(g) Show that

$$\lim_{\varepsilon \to 0^+} \int_{\varepsilon}^{1-\varepsilon} \cot \pi x e(-nx) \, dx = -i \int_0^1 (\cot \pi x) \sin 2\pi x \, dx .$$

(h) Conclude that

$$\lim_{\varepsilon \to 0^+} \int_{\varepsilon}^{1-\varepsilon} (\cot \pi x) e(-nx) \, dx = \begin{cases} -i & (n > 0), \\ 0 & (n = 0), \\ i & (n > 1). \end{cases}$$

18. Suppose that $f \in L^1(\mathbb{T})$, and recall the old-fashioned notation (3.2).

(a) Show that $(f * D_N)(x) = \frac{1}{2}a_0 + \sum_{n=1}^{N} (a_n \cos 2\pi n x + b_n \sin 2\pi n x)$.

(b) Show that $(f * D_N^*)(x) = \frac{1}{2}a_0 + \sum_{n=1}^{N-1} (a_n \cos 2\pi n x + b_n \sin 2\pi n x)$ $+ \frac{1}{2}(a_N \cos 2\pi N x + b_N \sin 2\pi N x)$.

(c) Show that $(f * \widetilde{D}_N)(x) = \sum_{n=1}^{N} (-b_n \cos 2\pi n x + a_n \sin 2\pi n x)$.

(d) Show that $(f * \Delta_N)(x) = \frac{1}{2}a_0 + \sum_{n=1}^{N} (1 - n/N)(a_n \cos 2\pi n x + b_n \sin 2\pi n x)$.

(e) Show that $(f * \widetilde{\Delta}_N)(x) = \sum_{n=1}^{N-1} (1 - n/N)(-b_n \cos 2\pi n x + a_n \sin 2\pi n x)$.

4.2. Special coefficients

We begin with a short review of summation by parts (aka partial summation), which is a discrete analogue of integration by parts.

Lemma 4.17. *For* $1 \le n \le N$, *let* c_n *and* e_n *be real or complex numbers, and let* $S_n = e_1 + e_2 + \cdots + e_n$. *Then*

$$\sum_{n=1}^{N} c_n e_n = \sum_{n=1}^{N-1} (c_n - c_{n+1}) S_n \ + c_N S_N \, . \tag{4.40}$$

Proof. Put $S_0 = 0$, so that $e_n = S_n - S_{n-1}$ for $1 \le n \le N$. Thus

$$\begin{aligned} \sum_{n=1}^{N} c_n e_n &= \sum_{n=1}^{N} c_n (S_n - S_{n-1}) \\ &= \sum_{n=1}^{N} c_n S_n - \sum_{n=1}^{N} c_n S_{n-1} \, . \end{aligned}$$

We reindex the second sum with $m = n - 1$ to see that the above is

$$= \sum_{n=1}^{N} c_n S_n - \sum_{m=0}^{N-1} c_{m+1} S_m \, .$$

For the indices running from 1 to $N - 1$ we combine the terms in the two sums. The term $c_N S_N$ for $n = N$ in the first sum is reported in (4.40). The term $c_1 S_0$ in the second sum makes no contribution since $S_0 = 0$. Thus we have the stated identity. □

If $0 < M < N$ and we apply the partial summation identity (4.40) first for M and then for N and subtract, we find that

$$\sum_{n=M+1}^{N} c_n e_n = \sum_{n=M}^{N-1} (c_n - c_{n+1}) S_n - c_M S_M + c_N S_N \, . \tag{4.41}$$

The two most familiar applications of partial summation are as follows.

Theorem 4.18. (Abel's Test) *If* $\sum_{n=1}^{\infty} e_n$ *is a convergent series, and if* $\{c_n\}$ *is a bounded monotonically decreasing sequence, then the series* $\sum_{n=1}^{\infty} c_n e_n$ *is convergent.*

Theorem 4.19. (Dirichlet's Test) *Let* $S_n = e_1 + e_2 + \cdots + e_n$. *If the sequence* $\{S_n\}$ *is bounded, and the sequence* $\{c_n\}$ *is monotonically decreasing to* 0, *then the series* $\sum_{n=1}^{\infty} c_n e_n$ *converges.*

Combined proof of both Tests. By the Cauchy principle for convergence, it suffices to show that the left hand side of (4.41) tends to 0 as $M, N \to \infty$ with $M < N$. In both tests the sequence S_n of partial sums is bounded; let S be chosen so that $|S_n| \le S$ for all n. By the triangle inequality, the sum on the right hand side of (4.41) has absolute value not exceeding

$$\sum_{n=M}^{N-1} (c_n - c_{n+1})S = S \sum_{n=M}^{N-1} (c_n - c_{n+1}) = S(c_M - c_N)\,.$$

By Theorem 0.3, a bounded monotonic sequence has a limit; let c denote the limit of the numbers c_n. Since $c_M \to c$ and $c_N \to c$, it follows that $c_M - c_N \to 0$ as $M, N \to \infty$. We note that

$$\begin{aligned} |c_N S_N - c_M S_M| &= |(c_N - c)S_N - (c_M - c)S_M + c(S_N - S_M)| \\ &\le |c_N - c|S + |c_M - c|S + |c||S_N - S_M|\,. \end{aligned}$$

The first two terms on the right hand side tend to 0 since c_n is tending to c and S is finite. It is only in treating the final term that we argue differently for the two tests. In the case of Abel's Test, the S_n tend to a limit, say $S_n \to s$. Then $S_N - S_M = (S_N - s) - (S_M - s) \to 0$, and $|c|$ is finite, so the term tends to 0. In Dirichlet's Test, $c = 0$, so the term is 0 for all M and N. □

Lemma 4.20. *Suppose that $c_0, c_1, \ldots$ are positive real numbers tending monotonically to 0. Then the series*

$$f(x) = \sum_{n=0}^{\infty} c_n e(nx) \tag{4.42}$$

converges for $0 < x < 1$, and converges uniformly for $\delta \le x \le 1 - \delta$ if $0 < \delta \le 1/2$.

Since each summand is continuous, the uniform convergence implies that $f(x)$ is continuous for $0 < x < 1$ (recall Theorem 0.5). However, $f(x)$ is not necessarily in $L^1(\mathbb{T})$, because it may be very large when x is near 0.

Proof. Let $S_0(x) = 0$, and for positive integers N put $S_N(x) = 1 + e(x) + \cdots + e((N-1)x)$. Then by (4.9) we know that $|S_N(x)| \le 1/\sin \pi x$ for $0 < x < 1$, and that $|S_N(x)| \le 1/\sin \pi\delta$ uniformly for $\delta \le x \le 1 - \delta$. By summing by parts we see that

$$\sum_{n=0}^{N} c_n e(nx) = \sum_{n=1}^{N-1} (c_{n-1} - c_n) S_n(x) + c_N S_N(x)\,. \tag{4.43}$$

Here the last term tends to 0 uniformly for $\delta \le x \le 1-\delta$, since $|S_{N+1}(x)| \le 1/\sin \pi\delta$ and $c_N \searrow 0$. The sum is uniformly convergent, since

$$\Big|\sum_{n=N+1}^{\infty} (c_{n-1} - c_n) S_n(x)\Big| \le \frac{1}{\sin \pi\delta} \sum_{n=N+1}^{\infty} (c_{n-1} - c_n) = \frac{c_N}{\sin \pi\delta} \to 0\,.$$

□

Although f is continuous in the open interval $(0, 1)$, it may be that f tends to infinity so rapidly as $x \to 0$ that $f \notin L^1(\mathbb{T})$. Before considering conditions that ensure that $f \in L^1(\mathbb{T})$, we show that if $f \in L^1(\mathbb{T})$, then the series (4.42) that defines f is indeed the Fourier Series of f.

Theorem 4.21. *Let f be defined as in (4.42) where $c_n \searrow 0$. If $f \in L^1(\mathbb{T})$, then $\widehat{f}(n) = c_n$ if $n \ge 0$, and $\widehat{f}(n) = 0$ if $n < 0$.*

Proof. Let k be an integer. We suppose first that x is not an integer. Then by partial summation we see that

$$\sum_{n=M+1}^{N} c_n e(nx)(1-e(-kx)) = (1-e(-kx))\sum_{M}^{N-1}(c_n - c_{n+1})S_n(x) + \big(1-e(-kx)\big)\big(c_N S_N(x) - c_M S_M(x)\big). \tag{4.44}$$

We know that $|S_n(x)| \le 1/|\sin \pi x|$. Also, $|\sin \pi x| \ge 2\|x\|$, so $|S_n(x)| \le 1/(2\|x\|)$. Now $|e(\delta) - 1| = |2\sin \pi\delta| \le 2\pi|\delta|$ for any real δ. Hence $|1-e(-kx)| \le 2\pi|k||x|$. Since $|1-e(-kx)|$ has period 1, we may suppose that $-1/2 \le x \le 1/2$. Thus we conclude that $|1-e(-kx)| \le 2\pi|k|\|x\|$. By these estimates we see that the absolute value of the right hand side above is

$$\begin{aligned} &\le \frac{|1-e(-kx)|}{|\sin \pi x|}(c_M - c_N) + \frac{|1-e(-kx)|}{|\sin \pi x|}(c_M + c_N) \\ &= \frac{2|1-e(-kx)|}{|\sin \pi x|}c_M \le \frac{2\pi|k|\|x\|}{\|x\|}c_M = 2\pi|k|c_M\,. \end{aligned} \tag{4.45}$$

Since $c_M \to 0$ as $M \to \infty$, this tends to 0 uniformly in x. If x is an integer, then both sides of (4.44) are 0, so the upper bound (4.45) holds in this case also. Since the series

$$\sum_{n=1}^{\infty} c_n e(nx)(1-e(-kx))$$

converges uniformly, we may integrate it term-by-term. Thus we find that

$$\begin{aligned} \widehat{f}(0) - \widehat{f}(k) &= \int_0^1 f(x)(1-e(-kx))\,dx = \sum_{n=1}^{\infty} c_n \int_0^1 e(nx)(1-e(-kx))\,dx \\ &= \begin{cases} c_0 - c_k & (k \ge 0), \\ c_0 & (k < 0). \end{cases} \end{aligned}$$

Now $c_k \to 0$ as $k \to \infty$ by hypothesis, and $\widehat{f}(k) \to 0$ as $k \to \infty$ by the Riemann–Lebesgue Lemma. Hence $\widehat{f}(0) = c_0$. On cancelling this from both sides we deduce also that $\widehat{f}(k) = c_k$ if $k > 0$, and that $\widehat{f}(k) = 0$ if $k < 0$. □

In the sum that defines $f(x)$ there is little to no cancelation among the first terms if x is near an integer. By partial summation we know that

$$\Big|\sum_{n=N+1}^{\infty} c_n e(nx)\Big| \le \frac{c_N}{\|x\|}\,.$$

Thus

$$|f(x)| \le c_0 + c_1 + \cdots + c_N + \frac{c_N}{\|x\|}\,.$$

If $\|x\| \ge 1/(N+1)$, then the above is

$$\le 2(c_0 + c_1 + \cdots + c_N) \tag{4.46}$$

since $c_n \ge c_N$ for $0 \le n \le N$. We use this bound for $1/(N+1) \le \|x\| \le 1/N$.

Theorem 4.22. *Suppose that f is defined by* (4.42) *where $c_n \searrow 0$. Then*

$$\|f\|_1 \le 2c_0 + 4\sum_{n=1}^{\infty} \frac{c_n}{n}.$$

Proof. Since $f(-x) = \overline{f(x)}$, it follows that $|f(-x)| = |f(x)|$. Hence by (4.46) we see that

$$\int_0^1 |f(x)|\,dx = 2\int_0^{1/2} |f(x)| = 2\sum_{n=2}^{\infty}\int_{1/(n+1)}^{1/n} |f(x)|\,dx$$

$$\le 4\sum_{N=2}^{\infty} \frac{(c_0 + c_1 + \cdots + c_N)}{N(N+1)}. \tag{4.47}$$

Now

$$\sum_{N=n}^{\infty} \frac{1}{N(N+1)} = \sum_{N=n}^{\infty}\Big(\frac{1}{N} - \frac{1}{N+1}\Big) = \frac{1}{n},$$

so the quantity (4.47) is

$$= 2c_0 + 2c_1 + 4\sum_{k=2}^{\infty}\frac{c_n}{n},$$

which gives the stated bound. □

For f defined as in (4.42), we can take the real part of f to obtain a cosine series and the imaginary part to obtain a sine series. Thus far we have treated the two equally, and we can write

$$f(x) = \sum_{n=0}^{\infty}(c_n - c_{n+1})S_n(x),$$

but when we take a closer look at the sums $S_n(x)$ we notice a difference in the distribution of the real part and imaginary parts. With a little calculation one finds that

$$S_n(x) = \frac{ie(-x/2)}{2\sin \pi x} - \frac{ie((n+1/2)x)}{2\sin \pi x}.$$

Thus for fixed x, the $S_n(x)$ lie on a circle with center $ie(-x/2)/(2\sin\pi x)$ and radius $1/(2|\sin \pi x|)$. Thus if x is small and positive, then

$$\operatorname{Re} S_n \doteqdot \frac{\sin 2\pi n x}{2\pi x}, \tag{4.48}$$

$$\operatorname{Im} S_n \doteqdot \frac{1-\cos 2\pi n x}{2\pi x}. \tag{4.49}$$

Since $\operatorname{Re} S_n(x)$ is oscillating sinusoidally with a mean near 0, this suggests that if the differences $c_n - c_{n+1}$ are monotonically decreasing, then we could obtain a better estimate by summing by parts a second time. On the other hand, $\operatorname{Im} S_n(x)$, though oscillating, has a large mean value, which suggests that the estimates that we have already derived are sharp. We now show that both of these first impressions are correct.

Theorem 4.23. *If*

$$f(x) = \sum_{n=1}^{\infty} b_n \sin 2\pi nx \tag{4.50}$$

where $b_n \searrow 0$, *then*

$$\frac{1}{\pi}\sum_{n=1}^{\infty}\frac{b_n}{n} \le \|f\|_1 \le 4\sum_{n=1}^{\infty}\frac{b_n}{n}\,. \tag{4.51}$$

It should be understood that in the above situation, $f \in L^1(\mathbb{T})$ if and only if $\sum_{n=1}^{\infty} b_n/n < \infty$.

Proof. For the lower bound, suppose that f is an odd function in $L^1(\mathbb{T})$. Then $\widehat{f}(0) = 0$, and from Theorem 3.7 we know that if $F(x) = \int_0^x f(u)\,du$, then $\widehat{F}(0) = \int_0^1 (1/2 - u)f(u)\,du$. Now $F(0) = 0$, so from Theorem 3.17 we see that

$$\widehat{F}(0) = -\lim_{N\to\infty}\sum_{\substack{n=-N\\ n\ne 0}}^{N}\frac{\widehat{f}(n)}{2\pi i n} = \frac{1}{2\pi}\sum_{n=1}^{\infty}\frac{b_n}{n}\,.$$

That is,

$$\sum_{n=1}^{\infty}\frac{b_n}{n} = \pi\int_0^1 (1-2u)f(u)\,du\,. \tag{4.52}$$

Note that this applies to any odd function $f \in L^1(\mathbb{T})$, without any condition on the b_n. Now $-1 \le 1-2u \le 1$ for $0 \le u \le 1$, so

$$\Big|\int_0^1 (1-2u)f(u)\,du\Big| \le \int_0^1 |(1-2u)f(u)|\,du \le \int_0^1 |f(u)|\,du = \|f\|_1\,.$$

Thus we have the lower bound in (4.51). The upper bound was already established in Theorem 4.22. ☐

For odd functions whose sine coefficients tend monotonically to 0, we can also characterize when the function is bounded and when it is continuous.

Theorem 4.24. *Suppose that* $b_n \searrow 0$, *and that* $f(x)$ *is defined as in* (4.50). *Then the following are equivalent:*

(a) $|f(x)|$ *is bounded*;

(b) nb_n *is a bounded sequence*;

(c) *The series* (4.50) *is boundedly convergent.*

Proof. (a) $\Longrightarrow$ (b). Let M be chosen so that $|f(x)| \le M$ for all x. As noted in (4.10), it follows that $|\sigma_N(x)| \le M$ for all x and all N. To bound b_n, we take $x = 1/(4n)$ and $N = 2n$. Thus

$$\sum_{k=1}^{2n}\Big(1-\frac{k}{2n}\Big)b_k \sin\frac{\pi k}{2n} \le M\,.$$

Now $\sin \pi x \ge 0$ for $0 \le x \le 1$, so the sine factor is nonnegative in every term. Thus the above sum is

$$\ge \sum_{k=1}^{n} \Big(1 - \frac{k}{2n}\Big) b_k \sin \frac{\pi k}{2n} \,.$$

Now $1 - k/(2n) \ge 1/2$ for $1 \le k \le n$, $b_k \ge b_n$, and $\sin \pi x \ge 2x$ for $0 \le x \le 1/2$, so the above is

$$\ge \frac{1}{2} b_n \sum_{k=1}^{n} \frac{k}{n} = (n+1) b_n / 4 \,.$$

Thus $nb_n \le 4M$ for all positive n.

(b) $\implies$ (c). Suppose that B is chosen so that $nb_n \le B$ for all positive n. That the series (4.50) converges for all x was established already in Lemma 4.20. In order to show that the series is boundedly convergent we need to establish a bound for

$$\sum_{n=M+1}^{N} b_n \sin 2\pi n x$$

that holds uniformly for all M, N, and x. Since f is an odd function with period 1, we may assume that $0 \le x \le 1/2$. We consider three cases.

Case 1. $Nx \le 2$. Since $\sin u \le u$ for all u, we see that

$$\sum_{n=M+1}^{N} b_n \sin 2\pi n x \le 2\pi x \sum_{n=M+1}^{N} n b_n \le 2\pi N x B \le 4\pi B \,.$$

Case 2. $Mx \ge 1$. Let $S_n(x) = \sum_{k=0}^{n-1} e(kx)$. By the summation by parts formula (4.41) we see that

$$\sum_{n=M+1}^{N} b_n e(nx) = \sum_{n=M}^{N-1} (b_n - b_{n+1}) S_n(x) \; - b_M S_M(x) + b_N S_N(x) \,.$$

From (4.9) we know that $|S_n(x)| \le 1/\sin \pi x$ for all n. Thus the above has absolute value not exceeding

$$\frac{1}{\sin \pi x} \sum_{n=M}^{N-1} (b_n - b_{n+1}) + \frac{b_M}{\sin \pi x} + \frac{b_N}{\sin \pi x} = \frac{2 b_M}{\sin \pi x} \,.$$

But $\sin \pi x \ge 2x$, so the above is

$$\le \frac{b_M}{x} = \frac{M b_M}{M x} \le B.$$

Case 3. $Mx < 1$, $Nx > 2$. Choose K so that $1 \le Kx \le 2$, and write

$$\sum_{n=M+1}^{N} b_n \sin 2\pi n x = \sum_{n=M+1}^{K} b_n \sin 2\pi n x + \sum_{n=K+1}^{N} b_n \sin 2\pi n x \,.$$

Then apply Case 1 to the first sum and Case 2 to the second sum. Thus

$$\Big| \sum_{n=M+1}^{N} b_n \sin 2\pi n x \Big| \le (6\pi + 1) B$$

in all cases.

(c) $\implies$ (a). This is trivial. We know that the series converges to $f(x)$ for all x. Since the partial sums are uniformly bounded, it follows that the limit $f(x)$ that they are tending to is also bounded. □

The same line of reasoning yields the following similar result.

Theorem 4.25. *Suppose that $b_n \searrow 0$, and that $f(x)$ is defined as in* (4.50). *Then the following are equivalent:*

(a) $|f(x)|$ *is continuous*;

(b) $nb_n \to 0$ *as* $n \to \infty$;

(c) *The series* (4.50) *is uniformly convergent.*

We now consider cosine series with convex coefficients.

Lemma 4.26. *Let $a_0, a_1, a_2, \ldots$ be a sequence of real numbers such that a_n tends to* 0 *as n tends to infinity, and $a_{n-1} - 2a_n + a_{n+1} \ge 0$ for all $n \ge 1$. Then*

(a) *The sequence a_n is monotonically decreasing*;

(b) $(a_n - a_{n+1})n \to 0$ *as* $n \to \infty$;

and

(c) $$\sum_{n=k}^{\infty}(a_{n-1} - 2a_n + a_{n+1})n = (a_{k-1} - a_k)k + a_k = ka_{k-1} - (k-1)a_k$$

for any positive integer k.

Proof. (a) Our second hypothesis concerning the a_n asserts that $a_{n-1} - a_n \ge a_n - a_{n+1}$. Thus $a_0 - a_1 \ge a_1 - a_2 \ge a_2 - a_3 \ge \cdots$, and these differences tend to 0 because the a_n tend to 0. Hence these differences are nonnegative, which is to say that $a_n \ge a_{n+1}$ for all $n \ge 0$.
(b) Suppose that $0 < M < N$. Then

$$a_M \ge a_M - a_N = \sum_{n=M+1}^{N}(a_{n-1} - a_n)\,.$$

The summands $a_{n-1} - a_n$ are decreasing, so they are all $\ge a_{N-1} - a_N$. Moreover, there are $N - M$ summands, so the sum above is

$$\ge (N - M)(a_{N-1} - a_N)\,.$$

Now a_M is small if M is large, and $N - M \ge N/2$ if $N \ge 2M$, so $N(a_{N-1} - a_N) \to 0$.
(c) Clearly

$$\sum_{n=k}^{N}(a_{n-1} - 2a_n + a_{n+1})n = \sum_{n=k}^{N}(a_{n-1} - a_n)n - \sum_{n=k}^{N}(a_n - a_{n+1})n\,.$$

After reindexing the second sum, we find that the above is

$$\begin{aligned}
&= \sum_{n=k}^{N}(a_{n-1} - a_n)n - \sum_{n=k+1}^{N+1}(a_{n-1} - a_n)(n-1) \\
&= (a_{k-1} - a_k)k - (a_N - a_{N+1})N + \sum_{n=k+1}^{N}(a_{n-1} - a_n) \\
&= (a_{k-1} - a_k)k - (a_N - a_{N+1})N + a_k - a_N\,.
\end{aligned}$$

By parts (a) and (b) this tends to $(a_{k-1} - a_k)k + a_k$ as $N \to \infty$. □

Theorem 4.27. *If $a_0, a_1, a_2, \ldots$ are real numbers such that $a_n \to 0$ as $n \to \infty$, and $a_{n-1} - 2a_n + a_{n+1} \ge 0$ for all $n \ge 1$, then*

(a)
$$\frac{1}{2}a_0 + \sum_{n=1}^{N} a_n \cos 2\pi x = \frac{1}{2}\sum_{n=1}^{N-1}(a_{n-1} - 2a_n + a_{n+1})n\Delta_n(x) + \frac{1}{2}(a_{N-1} - a_N)N\Delta_N(x) + \frac{1}{2}a_N D_N(x)$$

for all x;
(b) *the series*

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos 2\pi nx$$

converges for $0 < x < 1$;

(c)
$$f(x) = \frac{1}{2}\sum_{n=1}^{\infty}(a_{n-1} - 2a_n + a_{n+1})n\Delta_n(x)$$

for $0 < x < 1$;
(d) *$f(x) \ge 0$ for $0 < x < 1$;*

(e)
$$\int_0^1 f(x)\,dx = \frac{1}{2}a_0;$$

(f) *$\widehat{f}(\pm n) = \frac{1}{2}a_n$ for $n \ge 0$.*

The coefficients a_n may tend to 0 arbitrarily slowly. In Example 4.2 we saw that Fourier coefficients may tend to 0 arbitrarily slowly, but in that construction the Fourier Series was absolutely convergent, but most Fourier coefficients were 0. In the above we see that Fourier coefficients can tend to 0 arbitrarily slowly even when they are monotonic.

Consider the two series

$$f(x) = \sum_{n=1}^{\infty} \frac{\sin 2\pi nx}{\log(n+2)}, \qquad g(x) = \sum_{n=1}^{\infty} \frac{\cos 2\pi nx}{\log(n+2)}. \tag{4.53}$$

The coefficients here are monotonically decreasing to 0, so by taking real and imaginary parts in Lemma 4.20 we know that these series are convergent for $0 < x < 1$, and that f and g are continuous in this interval. Since

$$\sum_{n=1}^{\infty} \frac{1}{n\log(n+2)} = \infty,$$

we know by Theorem 4.23 that $f \notin L^1(\mathbb{T})$. By applying Theorem 4.27 to $\frac{1}{2\log 2}+g(x)$ we find that $g \in L^1(\mathbb{T})$. From the Riemann–Lebesgue Lemma we know that a sequence must tend to 0 at $\pm\infty$ in order to be a set of Fourier coefficients, but by this example we see that this condition by itself is not sufficient.

Put

$$G(x) = \sum_{n=2}^{\infty} \frac{\sin 2\pi nx}{n \log n}\,.$$

Since $g \in L^1(\mathbb{T})$, it follows by Theorem 3.17 that the Fourier Series of G converges to G. However, $G \notin A(\mathbb{T})$, so we see that Corollary 4.5 does not include Theorem 3.17, even when the coefficients are monotonic.

Proof. (a) We observe that

$$\begin{aligned}
\frac{1}{2}a_0 + \sum_{n=1}^{N} a_n \cos 2\pi nx &= \frac{1}{2}a_0 + \frac{1}{2}\sum_{n=1}^{N} a_n (D_n(x) - D_{n-1}(x)) \\
&= \frac{1}{2}a_0 + \frac{1}{2}\sum_{n=1}^{N} a_n D_n(x) - \frac{1}{2}\sum_{n=0}^{N-1} a_{n+1} D_n(x) \\
&= \frac{1}{2}\sum_{n=0}^{N-1} (a_n - a_{n-1}) D_n(x) \;+ \frac{1}{2}a_N D_N(x)\,. \qquad (4.54)
\end{aligned}$$

Here the sum over n is

$$\begin{aligned}
&\sum_{n=0}^{N-1} (a_n - a_{n+1})\big((n+1)\Delta_{n+1}(x) - n\Delta_n(x)\big) \\
&\quad = \sum_{n=0}^{N-1} (a_n - a_{n+1})(n+1)\Delta_{n+1}(x) - \sum_{n=0}^{N-1} (a_n - a_{n+1}) n\Delta_n(x) \\
&\quad = \sum_{n=1}^{N} (a_{n-1} - a_n) n\Delta_n(x) - \sum_{n=0}^{N-1} (a_n - a_{n+1}) n\Delta_n(x) \\
&\quad = \sum_{n=1}^{N-1} (a_{n-1} - 2a_n + a_{n+1}) n\Delta_n(x) + (a_{N-1} - a_N) N\Delta_N(x)\,.
\end{aligned}$$

On inserting this in (4.54), obtain the identity (a).
(b) By Lemma 4.26 (a) we know that the a_n are monotonically decreasing. Thus our series is the real part of the convergent series in Lemma 4.20.
(c) Suppose that $0 < x < 1$. Then the quantities $N\Delta_N(s)$ and $D_N(x)$ in (a) are bounded functions of N. Hence by Lemma 4.26 (a),(b) the last two terms in (a) tend to 0 as $N \to \infty$, and so we have the stated formula for $f(x)$.
(d) This follows from (c), since all terms are nonnegative.
(e) Integration and summation may be exchanged if the functions being summed are nonnegative, and $\int_0^1 \Delta_N(x)\,dx = 1$, so the stated result follows by taking $k = 1$ in Lemma 4.26 (c).
(f) Since $f(x) \geq 0$ and $\int_0^1 f(x)\,dx < \infty$, it follows by the principle of dominated

convergence (see Theorem 0.19) that for $k > 0$,

$$\widehat{f}(k) = \frac{1}{2}\sum_{n=1}^{\infty}(a_{n-1} - 2a_n + a_{n+1})n\widehat{\Delta_n}(k) = \frac{1}{2}\sum_{n=k+1}^{\infty}(a_{n-1} - 2a_n + a_{n+1})(n-k)\,.$$

This is the limit as $N \to \infty$ of

$$\begin{aligned}
&\frac{1}{2}\sum_{n=k+1}^{N}(a_{n-1} - 2a_n + a_{n+1})(n-k)\\
&\qquad = \frac{1}{2}\sum_{n=k+1}^{N}(a_{n-1} - a_n)(n-k) - \frac{1}{2}\sum_{n=k+1}^{N}(a_n - a_{n+1})(n-k)\\
&\qquad = \frac{1}{2}\sum_{n=k}^{N-1}(a_n - a_{n+1})(n+1-k) - \frac{1}{2}\sum_{n=k+1}^{N}(a_n - a_{n+1})(n-k)\\
&\qquad = \frac{1}{2}(a_k - a_{k+1}) - (a_N - a_{N+1})(N-k) + \frac{1}{2}\sum_{n=k+1}^{N-1}(a_n - a_{n+1})\\
&\qquad = \frac{1}{2}a_k - (a_N - a_{N+1})(N-k) - a_N.
\end{aligned}$$

By Lemma 4.26 (a),(b), the last two terms above tend to 0 as $N \to \infty$. Thus $\widehat{f}(k) = \frac{1}{2}a_k$ when $k > 0$. Since f is even, $\widehat{f}(-k) = \widehat{f}(k)$, so the proof is complete. □

4.3. Summability

Sometimes the existence of one limit implies the existence of another. We begin with a simple example of this phenomenon.

Theorem 4.28. *Suppose that $u_1, u_2, u_3, \ldots$ is an infinite sequence such that $\lim u_n = a$. Put $v_1 = u_1$, $v_2 = (u_1 + u_2)/2$, $v_3 = (u_1 + u_2 + u_3)/3$, and in general*

$$v_n = \frac{1}{n}\sum_{k=1}^{n} u_k\,.$$

Then $\lim_{n\to\infty} v_n = a$.

Proof. Put $U_n = u_n - a$, $V_n = v_n - a$. Then $\lim_{n\to\infty} U_n = 0$,

$$V_n = \frac{1}{n}\sum_{k=1} U_k,$$

and we wish to prove that $\lim_{n\to\infty} V_n = 0$. That is, it suffices to prove the theorem in the case $a = 0$.

Since $\lim_{n\to\infty} U_n = 0$, for any given $\varepsilon > 0$ there is an N such that $|U_n| < \varepsilon$ whenever $n > N$. By the triangle inequality,

$$|V_n| \le \frac{1}{n}\sum_{k=1}^{n}|U_k| = \frac{1}{n}\sum_{k=1}^{N}|U_k| + \frac{1}{n}\sum_{k=N+1}^{n}|U_k|\,.$$

Here the first sum is independent of n, and each summand in the second sum is $< \varepsilon$. Hence the above is

$$\leq \frac{1}{n} \sum_{k=1}^{N} |U_k| + \frac{n-N}{n} \varepsilon \,.$$

Here the second term is $< \varepsilon$, and the first term tends to 0 as $n \to \infty$. Thus $|V_n| < 2\varepsilon$ for all sufficiently large n. Since ε can be taken to be arbitrarily small, it follows that $\lim_{n\to\infty} V_n = 0$. □

Whether it is a good idea to average a sequence depends on the situation. If the sequence u_n is oscillatory and tends to a slowly, then the v_n may tend to a much more quickly. On the other hand, if u_n tends to a very quickly, then the sequence v_n of averages tends to a much more slowly. For example, if $u_n = 1/2^n$, then $v_n \asymp 1/n$.

Let s_n denote a partial sum of an infinite series, as in (4.1), and let σ_N be the average of the first N partial sums, as in (4.2). On taking $u_n = s_{n-1}$ and $v_n = \sigma_{n-1}$ we see from Theorem 4.28 that if the s_n tend to a limit, then the σ_N tend to the same limit. In other words, if a series converges to a, then it is also Cesàro summable to a. This has an immediate application to Fourier series, in view of Fejér's Theorem (Theorem 4.14):

Corollary 4.29. *Suppose that $f \in L^1(\mathbb{T})$. If f is continuous at x, and if the partial sums $s_N(x)$ tend to a limit, then that limit is $f(x)$.*

For power series we have the following fundamental result.

Theorem 4.30. (Abel) *If $\sum_{n=0}^{\infty} a_n$ converges, then*

$$\lim_{x\to 1^-} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n \,.$$

Proof. Let $s_n = \sum_{k=0}^{n} a_k$. We begin by observing that

$$(1-x) \sum_{n=0}^{\infty} s_n x^n = \sum_{n=0}^{\infty} a_n x^n \tag{4.55}$$

when $|x| < 1$. To see why this is so, observe that the left hand side is

$$\sum_{n=0}^{\infty} s_n x^n - \sum_{n=0}^{\infty} s_n x^{n+1} = \sum_{n=0}^{\infty} s_n x^n - \sum_{n=1}^{\infty} s_{n-1} x^n = s_0 + \sum_{n=1}^{\infty} (s_n - s_{n-1}) x^n \,.$$

But $s_n - s_{n-1} = a_n$ for $n > 0$, so we have (4.55). Let $S = \sum_{n=0}^{\infty} a_n$, and put $S_n = s_n - S$. Since

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \tag{4.56}$$

for $|x| < 1$, it follows that

$$(1-x)\Big(\sum_{n=0}^{\infty} s_n x^n\Big) - S = (1-x)\Big(\sum_{n=0}^{\infty} s_n x^n\Big) - (1-x)\sum_{n=0}^{\infty} S x^n = (1-x)\sum_{n=0}^{\infty} S_n x^n \,.$$

Our task now is to show that if the S_n tend to 0, then the right hand side above also tends to 0. That is, we have reduced our problem to the special case $S = 0$. Let $\varepsilon > 0$ be given, and let N be so large that $|S_n| < \varepsilon$ for all $n \ge N$. Write the right hand side above as

$$(1-x)\sum_{n=0}^{N-1} S_n x^n + (1-x)\sum_{n=N}^{\infty} S_n x^n = T_1 + T_2,$$

say. Now

$$|T_1| \le (1-x)\sum_{n=0}^{N-1} |S_n| \tag{4.57}$$

for $0 \le x \le 1$, and

$$|T_2| \le (1-x)\sum_{n=N}^{\infty} \varepsilon x^n \le \varepsilon(1-x)\sum_{n=0}^{\infty} x^n = \varepsilon$$

by (4.56). From (4.57) we see that $T_1 \to 0$ as $x \to 1^-$. Thus there is a $\delta > 0$ such that $|T_1| < \varepsilon$ when $1 - \delta < x < 1$. Hence

$$\left|(1-x)\sum_{n=0}^{\infty} S_n x^n\right| < 2\varepsilon$$

for $1 - \delta < x < 1$. Since ε can be taken arbitrarily small, it follows that the limit is 0 as $x \to 1^-$, and the proof is complete. □

Although it was a simple matter to verify (4.55), there is a different approach that is also instructive. By dividing both sides by $1 - x$ and using (4.56), we see that (4.55) is equivalent to the identity

$$\Big(\sum_{j=0}^{\infty} x^j\Big)\Big(\sum_{k=0}^{\infty} a_k x^k\Big) = \sum_{n=0}^{\infty} s_n x^n\,.$$

Here the sums on the left hand side are absolutely convergent for $|x| < 1$, and hence can be arbitrarily rearranged. The general term is $a_k x^{j+k}$. We choose an n, and consider those pairs j, k for which $j + k = n$. Thus we see that the coefficient of x^n is $a_0 + a_1 + \cdots + a_n = s_n$. Thus we have the identity above, and hence (4.55).

Corollary 4.31. *If* $\sum_{n=0}^{\infty} a_n e(n\theta)$ *converges, then*

$$\lim_{r\to 1^-}\sum_{n=0}^{\infty} a_n r^n e(n\theta) = \sum_{n=0}^{\infty} a_n e(n\theta)\,.$$

If a power series $f(z)$ has radius of convergence $R > 1$, then f is continuous in the closed disk $|z| \le 1$, which means that the above is trivial in this case. Thus the main interest in Corollary 4.31 is when f has radius of convergence 1. One should note that in Corollary 4.31 we take only a radial limit, which is less than asserting that $f(z)$ is continuous at $e(\theta)$. Our method of proof could be developed to show that $f(z)$ tends to $f(e(\theta))$ as z tends to $e(\theta)$ within a sector, as depicted in Figure 4.3 (a). However, examples can be constructed to show that the limit need not exist if z approaches $e(\theta)$ along a path that approaches $|z| = 1$ tangentially, as in Figure 4.3 (b).

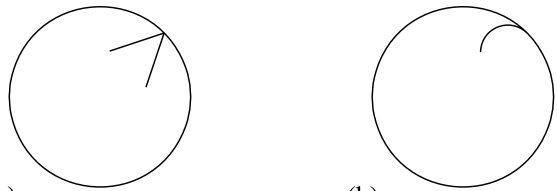


Figure 4.3. (a) A sector that bounds the approach; (b) A tangential approach.

One may note that the proof of Theorem 4.30 bears a striking resemblance to the proof of Theorem 4.28. The similarity is more than just superficial. In Theorem 4.28 we consider averages of the u_n. Put $w_n = (1-x)x^n$. Then the left hand side of (4.55) is $\sum_{n=0}^\infty w_n s_n$, which is a weighted average of the s_n, in the sense that $w_n \ge 0$ for all n, and $\sum_{n=0}^\infty w_n = 1$ by (0.3). Moreover, when x is just slightly less than 1, the initial weights $w_0, w_1, \ldots$ are very small but approximately equal. Thus having x near 1 in the proof of Theorem 4.30 is analogous to having n large in the proof of Theorem 4.28.

If $\lim_{x\to 1^-} \sum_{n=0}^\infty a_n x^n$ exists and has the value a, then we say that the series $\sum a_n$ is *Abel-summable* to a, and we write

$$\sum_{n=0}^\infty a_n = a \quad \text{(A)}.$$

Example 4.3. Let $a_n = (-1)^n$. Then

$$\sum_{n=0}^\infty a_n x^n = \sum_{n=0}^\infty (-1)^n x^n = \sum_{n=0}^\infty (-x)^n = \frac{1}{1+x}$$

for $|x| < 1$, by (4.56). On letting x tend to 1 from below, we find that

$$\sum_{n=0}^\infty (-1)^n = 1/2 \quad \text{(A)}.$$

Theorems 4.28 and 4.30 are called *abelian* in honor of Abel, who proved Theorem 4.30. In general, an abelian theorem is one in which one limit implies another. Usually, the second limit is some sort of weighted average of the first. From Theorem 4.28 we saw that if a series is convergent, then it is Cesàro-summable to the same value. In Example 4.1 we found that the converse is false, and in Example 4.3 we found that the converse of Abel's Theorem (Theorem 4.30) is also false. While converses of abelian theorems are usually false, often one can prove a partial converse by imposing an extra hypothesis. Such partial converses are called *tauberian*, after Tauber, who proved a partial converse of Abel's Theorem. The following is a good first example of a tauberian theorem.

Theorem 4.32. (Hardy) *Suppose that the series* $\sum_{n=1}^\infty a_n$ *is Cesàro-summable to* a *and that there is a constant* $C > 0$ *such that*

$$|a_n| \le \frac{C}{n} \tag{4.58}$$

for all n. *Then* $\sum_{n=1}^\infty a_n$ *converges to* a.

Here (4.58) is the new hypothesis, called the *tauberian hypothesis*, which permits the converse. Suppose the $f(n)$ is positive and increasing to $+\infty$. No matter how slowly f increases, one can construct examples to show that if (4.58) were replaced by the weaker hypothesis

$$|a_n| \le \frac{f(n)}{n},$$

then Theorem 4.32 would become false. That is, the hypothesis (4.58) is the weakest hypothesis of its kind that enables the converse implication.

Proof. Let $s_n = \sum_{k=1}^n a_k$ and $\sigma_n = \sum_{j=1}^n \left(1 - \frac{j-1}{n}\right)a_j$. We have to show that if $\sigma_n \to a$ as $n \to \infty$, then also $s_n \to a$. We note that

$$(n+h)\sigma_{n+h} - n\sigma_n = \sum_{j=1}^{n+h}(n+h+1-j)a_j \;-\; \sum_{j=1}^{n}(n+1-j)a_j\,.$$

For $1 \le j \le n$, the difference between the coefficients of a_j is simply h. Thus the above is

$$= hs_n + \sum_{j=n+1}^{n+h}(n+h+1-j)a_j$$

We subtract ha from both sides, and divide through by h. Thus the above can be written

$$s_n - a = \frac{n+h}{h}(\sigma_{n+h} - a) - \frac{n}{h}(\sigma_n - a) - \sum_{j=n+1}^{n+h}\frac{n+h+1-j}{h}a_j\,.$$

In the remaining sum, $|a_j| \le C/j \le C/n$, the coefficient of a_j is at most 1, and there are h terms. Hence by the triangle inequality,

$$|s_n - a| \le \frac{n+h}{h}|\sigma_{n+h} - a| + \frac{n}{h}|\sigma_n - a| + \frac{Ch}{n}\,. \tag{4.59}$$

Choose an $\varepsilon > 0$, take $h = [\varepsilon n]$ and let n tend to infinity. The ratio $(n+h)h$ is bounded, and $\sigma_{n+h} - a$ tends to 0, so the first term on the right hand side tends to 0. The ratio n/h is also bounded, and $\sigma_n - a$ tends to 0, so the second term also tends to 0. Finally, the last term is $\le C\varepsilon$. That is,

$$|s_n - a| \le (C+1)\varepsilon$$

for all sufficiently large n. But ε can be arbitrarily small, so we conclude that $s_n \to a$ as $n \to \infty$. □

We might say that the σ_n is obtained from the s_n by smoothing. In the proof just completed, we recovered the s_n from the σ_n by an “unsmoothing” that hinges on the identity

$$\max(n+h-j,0) - \max(n-j,0) = \begin{cases} h & (0 \le j \le n), \\ n+h-j & (n \le j \le n+h), \\ 0 & (j \ge n+h). \end{cases} \tag{4.60}$$

A graphical presentation of this identity is given in Figure 4.4.

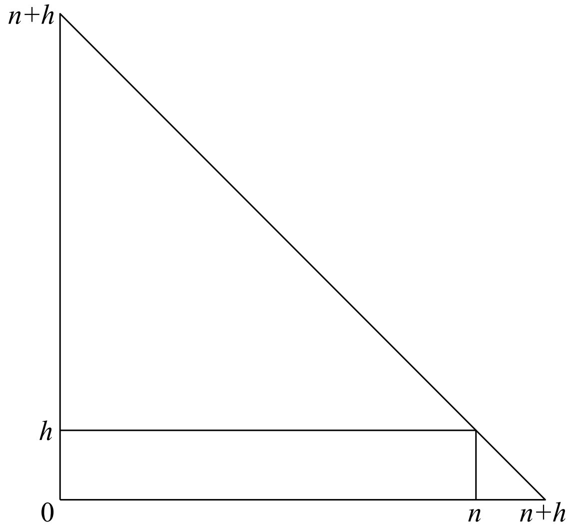


Figure 4.4. Graphical depiction of the unsmoothing identity (4.60).

On combining Fejér's theorem (Theorem 4.14) with Hardy's tauberian theorem (Theorem 4.32) we obtain

Corollary 4.33. *Suppose that $f \in L^1(\mathbb{T})$ and that $|n\widehat{f}(n)|$ is bounded. Then the partial sums $s_N(x)$ of the Fourier series of f converge to $\frac{1}{2}\big(f(x^-)+f(x^+)\big)$ whenever the one-sided limits $f(x^-)$ and $f(x^+)$ both exist and are finite.*

In the context of Hardy's tauberian theorem, suppose that a sequence $\varepsilon_n \searrow 0$ has the property that $|\sigma_n - a| \le \varepsilon_n$ for all n. Then from the key inequality (4.59) we find that

$$|s_n - a| \le \frac{3n}{n}\varepsilon_n + \frac{Ch}{n}\,.$$

On taking $h = [\sqrt{\varepsilon_n}n]$ we deduce that $|s_n - a| \le C_1\sqrt{\varepsilon_n}$. Thus a quantitative form of the hypothesis that $\sigma_n \to a$ yields a corresponding quantitative form of the conclusion that $s_n \to a$. This observation is particularly valuable when we apply Hardy's theorem to a Fourier Series. If $|n\widehat{f}(n)| \le C$ for all n, then with $a_n = \widehat{f}(n)e(nx) + \widehat{f}(-n)e(-nx)$ we have $|na_n| \le 2C$ *uniformly* in x. Thus if $\sigma_N(x) \to f(x)$ uniformly for x in a certain set, then $s_N \to f(x)$ uniformly for all x in that same set. Of course this applies only to those Fourier Series for which $|n\widehat{f}(n)|$ is a bounded sequence.

Abel summability involves considerably more smoothing than Cesàro summability, but we can still unsmooth, as follows.

Theorem 4.34. (Littlewood) *If $\sum_{n=0}^{\infty} a_n = a$ (A), and (4.58) holds for all n, then the series $\sum_{n=0}^{\infty} a_n$ converges to a.*

It is remarkable that when unsmoothing the abelian weights, the tauberian hypothesis needed is still the same one as in the Cesàro case.

In the argument below we employ a useful notational device: If S is a set, then the *characteristic function* of S, denoted χ_S, is

(4.61) $$\chi_S(x) = \begin{cases} 1 & (\text{if } x \in S), \\ 0 & (\text{otherwise}). \end{cases}$$

This convention is observed throughout mathematics, but in probability theory this function is called the *indicator function*, denoted I_S, and what a probabilist calls a characteristic function is what we call a Fourier transform.

Proof. Suppose that $\varepsilon > 0$ is given. Let $\mathcal{J} = [1/e, 1]$ and $\mathcal{K} = [e^{-1-\varepsilon}, e^{-1+\varepsilon}]$. Assume for the moment that there exists a polynomial $P(x)$ such that

(4.62) $$|P(x) - \chi_{\mathcal{J}}(x)| \le \varepsilon x(1-x) + 5\chi_{\mathcal{K}}(x)$$

for $0 \le x \le 1$. We note that

(4.63) $$s_N = \sum_{n=0}^{N} a_n = \sum_{n=0}^{\infty} a_n \chi_{\mathcal{J}}\big(e^{-n/N}\big)\,.$$

From (4.62) we see that $P(0) = 0$ and that $P(1) = 1$. Thus P is of the form $P(x) = \sum_{r=1}^{R} c_r x^r$ with $\sum_{r=1}^{R} c_r = 1$. If on the right hand side above we replace $\chi_{\mathcal{J}}$ by P, the expression becomes

(4.64) $$\sum_{n=0}^{\infty} a_n P\big(e^{-n/N}\big) = \sum_{n=0}^{\infty} a_n \sum_{r=1}^{R} c_r e^{-rn/N} = \sum_{r=1}^{R} c_r \sum_{n=0}^{\infty} a_n e^{-rn/N} = \sum_{r=1}^{R} c_r f\big(e^{-r/N}\big)$$

where $f(x) = \sum_{n=0}^{\infty} a_n x^n$. But $f(e^{-r/N}) \to a$ as $N \to \infty$, and $\sum_{r=1}^{R} c_r = 1$, so

(4.65) $$\lim_{N\to\infty} \sum_{r=1}^{R} c_r f\big(e^{-r/N}\big) = a\,.$$

By subtracting (4.64) from (4.63) and applying the triangle inequality we see that

(4.66) $$\Big| s_N - \sum_{r=1}^{R} c_r f\big(e^{-r/N}\big)\Big| \le \sum_{n=0}^{\infty} |a_n| \big|P(e^{-n/N}) - \chi_{\mathcal{J}}(e^{-n/N})\big|$$

which by (4.62) is

(4.67) $$\le \varepsilon \sum_{n=0}^{\infty} |a_n| e^{-n/N}\big(1 - e^{-n/N}\big) \; + \; 5 \sum_{(1-\varepsilon)N \le n \le (1+\varepsilon)N} |a_n|$$

(4.68) $$= T_1 + T_2,$$

say. If $0 \le u \le 1$, then $1 - e^{-u} \asymp u$. Thus

$$\sum_{n=0}^{N} |a_n| e^{-n/N}\big(1 - e^{-n/N}\big) \; \le \; \sum_{n=0}^{N} |a_n| \big(1 - e^{-n/N}\big) \; \le \; \frac{C_1}{N} \sum_{n=0}^{N} n|a_n|\,.$$

Here the term $n = 0$ makes no contribution, and $n|a_n|$ is bounded for $n > 0$ by (4.58), so the above is $\le C_2$. To estimate $\sum_{n>N} |a_n| e^{-n/N}$ we consider dyadic blocks. Suppose that $2^k N < n \le 2^{k+1} N$. There are $2^k N$ terms in this block, and

$|a_n| \le C/(2^k N)$ for all n in this range, so the contribution is $\le Ce^{-2^k}$. The sum of this over k converges, so $T_1 < C_3\varepsilon$. As for T_2, we note that there are $\asymp \varepsilon N$ summands, each of which is $< C_4/N$, so $T_2 < C_5\varepsilon$. On combining this with (4.65), we deduce that

$$|s_N - a| < C_6\varepsilon$$

for all sufficiently large N. Since ε may be taken arbitrarily small, it follows that $s_N \to a$ as $N \to \infty$.

It remains to show that there is a polynomial P with the property (4.62). Let $g(x) = -1/(1-x)$ for $0 \le x \le e^{-1-\varepsilon}$, let $g(x) = 1/x$ for $e^{-1+\varepsilon} \le x \le 1$, and in the interval $(e^{-1-\varepsilon}, e^{-1+\varepsilon})$ let g be linear, interpolating between $g(e^{-1-\varepsilon})$ and $g(e^{-1+\varepsilon})$, so that g is continuous on $[0,1]$. By the Weierstrass approximation theorem (Corollary 4.6) there is a polynomial $Q(x)$ such that $|Q(x) - g(x)| < \varepsilon$ for all $x \in [0,1]$. Now put

$$h(x) = \frac{\chi_{\mathcal{J}}(x) - x}{x(1-x)} = \begin{cases} -1/(1-x) & (0 \le x < 1/e), \\ 1/x & 1/e \le x \le 1). \end{cases} \tag{4.69}$$

Thus h is the same as g except that in the interval $\mathcal{K}$, h has a jump discontinuity while g is continuous. Thus $|Q(x) - h(x)| < 5$ for $x \in \mathcal{K}$ (assuming that ε is small), and $|Q(x) - h(x)| < \varepsilon$ otherwise. On rewriting (4.69), we see that

$$\chi_{\mathcal{J}}(x) = x + x(1-x)h(x)\,.$$

If we replace h by Q, then we obtain a polynomial $P(x) = x + x(1-x)Q(x)$ that has the property (4.62). □

Exercises

1. Put $s_N = \sum_{n=1}^N a_n$, and suppose that

$$\lim_{N\to\infty} \frac{s_N}{N} = a\,. \tag{4.70}$$

(a) Put $S_N = s_N - aN = \sum_{n=1}^N (a_n - a)$. Show that (4.70) is equivalent to

$$\lim_{N\to\infty} \frac{S_N}{N} = 0\,. \tag{4.71}$$

(b) Let $\sigma_N = \frac{1}{N}\sum_{n=1}^N s_n$. Show that

$$\sigma_N = \frac{1}{N}\sum_{k=1}^N (N+1-k)a_k\,.$$

Explain why

$$\frac{1}{N}\sum_{n=1}^N S_n = \sigma_N - \frac{1}{2}(N+1)a\,.$$

(c) Show that

$$\lim_{N\to\infty} \frac{2}{N+1}\sum_{n=1}^N S_n = 0\,.$$

(d) Conclude that (4.70) implies that

$$\lim_{N\to\infty} \frac{2}{N(N+1)} \sum_{n=1}^{N} (N+1-n)a_n = a\,. \tag{4.72}$$

2. (Frobenius) The object of this exercise is to prove that

$$\sum_{n=0}^{\infty} a_n = a \quad (\mathrm{C}) \qquad \Longrightarrow \qquad \sum_{n=0}^{\infty} a_n \quad (\mathrm{A})\,.$$

Thus it is no coincidence that Examples 4.1 and 4.3 produced the same values.

(a) Let

$$s_n = \sum_{k=0}^{n} a_k, \qquad \sigma_n = \frac{1}{n}\sum_{k=0}^{n-1} s_k\,.$$

Show that

$$\sum_{n=0}^{\infty} s_n x^n = (1-x)\sum_{n=0}^{\infty}(n+1)\sigma_{n+1}x^n\,.$$

(b) Deduce that

$$\sum_{k=0}^{\infty} a_k x^k = (1-x)^2 \sum_{n=0}^{\infty}(n+1)\sigma_{n+1}x^n$$

for $|x| < 1$.

(c) Show that

$$\sum_{n=0}^{\infty}(n+1)x^n = \frac{1}{(1-x)^2}$$

for $|x| < 1$.

(d) Deduce that

$$\sum_{k=0}^{\infty} a_k x^k \ - a = (1-x)^2 \sum_{n=0}^{\infty}(n+1)(\sigma_{n+1}-a)x^n\,.$$

(e) Suppose that $|\sigma_n - a| < \varepsilon$ for all $n > N$, and write the right hand side above as

$$(1-x)^2 \sum_{n=0}^{N-1}(n+1)(\sigma_{n+1}-a)x^n + (1-x)^2 \sum_{n=N}^{\infty}(n+1)(\sigma_{n+1}-a)x^n = T_1 + T_2\,.$$

(f) Explain why

$$|T_1| \le (1-x)^2 \sum_{n=0}^{N-1}(n+1)|\sigma_{n+1}-a|$$

uniformly for $0 \le x \le 1$.

(g) Show that

$$|T_2| \le \varepsilon$$

for all x, $0 \le x < 1$.

(h) Conclude that $\sum a_n = a$ (A).

3. In this Exercise, we derive a variant of Theorem 4.32. We assume that $\sum_{n=1}^{\infty} a_n = a$ (C). Instead of the tauberian hypothesis (4.58), we assume that

$$a_n \geq 0 \tag{4.73}$$

for all n. Let s_n and σ_n be defined as in the proof of Theorem 4.32.

(a) Explain why

$$n\sigma_n - (n-h)\sigma_{n-h} \leq hs_n \leq (n+h)\sigma_{n+h} - n\sigma_n \,.$$

(b) Deduce that

$$\frac{n(\sigma_n - a) - (n-h)(\sigma_{n-h} - a)}{h} \leq s_n - a \leq \frac{(n+h)(\sigma_{n+h} - a) - n(\sigma_n - a)}{n} \,.$$

(c) Take $h = [\varepsilon n]$, and let n tend to infinity. Conclude that $s_n \to a$ as $n \to \infty$. That is, the series $\sum_{n=1}^{\infty} a_n$ converges to a.

4. In this exercise, we derive a further variant of Theorem 4.32. We assume that $\sum_{n=1}^{\infty} a_n = a$ (C). In place of the tauberian hypothesis (4.73) we have a weaker hypothesis, namely that there is a constant $C > 0$ such that

$$a_n \geq -C/n \tag{4.74}$$

for all n.

(a) Explain why

$$n\sigma_n - (n-h)\sigma_{n-h} - \frac{Ch^2}{n} - h \leq hs_n \leq (n+h)\sigma_{n+h} - n\sigma_n + \frac{Ch^2}{n} \,.$$

(b) Deduce that

$$\frac{n(\sigma_n - a) - (n-h)(\sigma_{n-h} - a)}{h} - \frac{Ch}{n-h} \leq s_n - a \leq \frac{(n+h)(\sigma_{n+h} - a) - n(\sigma_n - a)}{h} + \frac{Ch}{n} \,.$$

(c) Take $h = [\varepsilon n]$, and let n tend to infinity. Deduce that $\sum_{n=1}^{\infty} a_n$ converges to a.

(d) Explain why this result implies the result of the preceding exercise.

(e) Explain why this result implies Theorem 4.32. (Caution: The a_n in Theorem 21 may be complex, so it is necessary to separate the real and imaginary parts, and treat them separately.)

5. We now establish a tauberian companion to the abelian result of Exercise 1. We suppose that (4.72) holds, and our tauberian hypothesis is that there is a constant $C > 0$ such that

$$|a_n| \leq C \tag{4.75}$$

for all n.

(a) Explain why

$$\sum_{n=1}^{N+H}(N+H+1-n)(a_n-a)-\sum_{n=1}^{N}(N+1-n)(a_n-a)$$
$$=H(s_N-aN)+\sum_{n=N+1}^{N+H}(N+H+1-n)(a_n-a)\,.$$

(b) Deduce that

$$\begin{aligned}\frac{s_N}{N}-a&=\frac{1}{NH}\sum_{n=1}^{N+H}(N+H+1-n)(a_n-a)-\frac{1}{NH}\sum_{n=1}^{N}(N+1-n)(a_n-a)\\&\quad-\frac{1}{NH}\sum_{n=N+1}^{N+H}(N+H+1-n)(a_n-a)\\&=T_1+T_2+T_3\,.\end{aligned}$$

(c) Show that $|a|\le C$.
(d) Deduce that $|a_n-a|\le 2C$ for all n.
(e) Take $H=[\varepsilon N]$ and let $N\to\infty$. Show that $T_1\to 0$ and that $T_2\to 0$.
(f) Show that $|T_3|\le 2CH/N$.
(g) Deduce that $s_N/N\to a$.

6. Let f be a function with a continuous second derivative.
(a) Explain why

$$f(x+h)=f(x)+hf'(x)+\int_x^{x+h}(x+h-u)f''(u)\,du\,.$$

(b) Deduce that

$$f'(x)=\frac{f(x+h)}{h}-\frac{f(x)}{h}-\int_0^h(1-v/h)f''(x+v)\,dv=T_1+T_2+T_3\,.$$

(c) Suppose that $|f(x)|\le C_0$ for all $x\ge X$, and that $|f''(x)|\le C_2$ for all $x\ge X$. Show that $|T_1|\le C_0/h$, that $|T_2|\le C_0/h$, and that $|T_3|\le C_2h/2$.
(d) By making a suitable choice of h, conclude that $|f'(x)|\le 2\sqrt{C_0C_2}$ for all $x\ge X$.
(e) Suppose that $f\to 0$ as $x\to\infty$, and that f'' is bounded. Deduce that $f'\to 0$ as $x\to\infty$.

7. Suppose that $\int_0^\infty g(x)\,dx$ converges, and that $g'(x)$ is bounded. Deduce that $g(x)\to 0$ as $x\to\infty$. (Hint: Take $f(x)=\int_x^\infty g(u)\,du$ in the preceding exercise.)

4.4. Summability kernels

We say that a sequence $K_N(x)$ of functions in $L^1(\mathbb{T})$ constitute a *summability kernel* if

$$\int_0^1 K_N(x)\,dx=1 \tag{4.76}$$

for all N, if there is a constant $C > 0$ such that

$$\int_0^1 |K_N(x)|\,dx \le C \tag{4.77}$$

for all N, and if

$$\lim_{N\to\infty} \int_\delta^{1-\delta} |K_N(x)|\,dx = 0\,. \tag{4.78}$$

In the case of the Fejér kernel, (4.76) is given by (4.7), (4.77) holds with $C = 1$ because by (4.8) the Fejér kernel is nonnegative, and (4.78) follows from (4.14). The Fejér kernel has additional helpful properties, but it is the three properties (4.76)–(4.78) that enable us to prove the important uniform approximation theorem of continuous functions (Theorem 4.1), and the convergence in norm for arbitrary $f \in L^1(\mathbb{T})$ (Theorem 4.7). Thus these theorems can be generalized to other summability kernels. The Dirichlet kernel satisfies (4.76), in view of (3.46), but

$$\int_0^1 |D_N(x)|\,dx \asymp \log N,$$

and

$$\int_\delta^{1-\delta} |D_N(x)|\,dx \asymp \log \frac{1}{\delta},$$

so both (4.77) and (4.78) fail for the Dirichlet kernel. Theorems 4.1, 4.7, 4.14, 4.15 all fail when the Cesàro partial sum $\sigma_N(x)$ is replaced by the unweighted partial sum $s_N(x)$. In each case, the reason for the failure can be traced back to the fact that $\int_0^1 |D_N(x)|\,dx$ is unbounded.

A summability kernel is not necessarily nonnegative. The inequalities (4.10) and (4.11) follow from from the nonnegativity of the Fejér kernel, and similar inequalities would apply to $(K_N * f)(x)$ if $K_N(x) \ge 0$ for all x and (4.76) holds. Fejér's pointwise convergence theorem (Theorem 4.14) also does not extend to arbitrary summability kernels, as its proof depends not only on (4.77) and (4.78) but also on a uniform decay property:

$$\lim_{N\to\infty} \max_{\delta\le x\le 1-\delta} |K_N(x)| = 0,$$

which for the Fejér kernel is given by (4.12).

Exercises

1. Using the proof of Theorem 4.1 as a model, give a detailed proof that if $f \in C(\mathbb{T})$ and K_N is a summability kernel,

   $$|f(x) - (f * K_N)(x)| < \varepsilon$$

   holds uniformly in x, if $N > N_0(\varepsilon)$.

2. Using the proof of Theorem 4.7 as a model, give a detailed proof that if $f \in L^1(\mathbb{T})$ and K_N is a summability kernel, then

   $$\lim_{N\to\infty} \int_0^1 |f(x) - (f * K_N)(x)|\,dx = 0\,.$$

3. For $0 \le r < 1$, the *Poisson kernel* is

$$P_r(x) = \sum_{n=-\infty}^{\infty} r^{|n|} e(nx) = 1 + 2\sum_{n=1}^{\infty} r^n \cos 2\pi n x \,. \tag{4.79}$$

In this context, $r \to 1^-$ corresponds to $N \to \infty$ for discretely indexed kernels.

(a) Let r be fixed, $0 \le r < 1$. Show the series defining $P_r(x)$ is absolutely convergent, that $P_r(x)$ is a continuous function of x, and that $\widehat{P_r}(n) = r^{|n|}$ for all n.

(b) By two applications of the formula for the sum of a geometric series, show that
$$P_r(x) = \frac{re(x)}{1-re(x)} + \frac{1}{1-re(-x)} \,.$$

(c) Deduce that
$$P_r(x) = \frac{1-r^2}{1-2r\cos 2\pi x + r^2} \,.$$

(d) Show that
$$P_r(x) = \frac{1-r^2}{(1-r)^2 + 4r\sin^2 \pi x} \,.$$

(e) Show that the denominator above is always $\ge (1-r)^2$.

(f) Show that for fixed r, $P_r(x)$ is a decreasing function of x for $0 \le x \le 1/2$.

(g) Show that $\int_0^1 P_r(x)\,dx = 1$.

(h) Show that $P_r(x) \ge 0$ for all x.

(i) Show that if $1/2 \le r < 1$, then
$$P_r(x) \le \frac{1-r}{\sin^2 \pi x}$$
uniformly in x.

(j) Deduce that P_r is a summability kernel. Show that if $f \in L^1(\mathbb{T})$, then
$$(f * P_r)(x) = \sum_{n=-\infty}^{\infty} r^{|n|} \widehat{f}(n) e(nx) \,.$$

(k) Show that $P_r(0) \sim 2/(1-r)$ as $r \to 1^-$.

(l) $P_r(x)$ has a peak at $x = 0$. About how wide is this peak?

4. Suppose that $f \in L^1(\mathbb{T})$, that f is odd, and that $f(x) \ge 0$ for $0 \le x \le 1/2$.

(a) Show that
$$(f * P_r)(x) = \int_0^{1/2} f(u)\big(P_r(x-u) - P_r(x+u)\big)\,du \,.$$

(b) Show that if $0 \le x \le 1/2$ and $0 \le u \le 1/2$, then $P_r(x-u) \ge P_r(x+u)$.

(c) Conclude that $(f * P_r)(x) \ge 0$ for $0 \le x \le 1/2$.

5. (a) Show that $x^n - y^n = (x-y)(x^{n-1} + x^{n-2}y + \cdots + y^{n-1})$.

(b) Deduce that
$$\sin 2\pi n x = (\sin 2\pi x)(e((n-1)x) + e((n-3)x) + \cdots + e(-(n-1)x)) \,.$$

(c) Deduce that $|\sin 2\pi n x| \le n |\sin 2\pi x|$ for all x.

(d) From now on, assume that f is an odd function in $L^1(\mathbb{T})$ such that $f(x) \geq 0$ for $0 \leq x \leq 1/2$. In the formula

$$b_1 = 2\int_0^1 f(x)\sin 2\pi x\,dx,$$

explain why the integrand is everywhere nonnegative.

(e) Deduce that $b_1 \geq 0$.

(f) Show that

$$\int_0^1 (n\sin 2\pi x \pm \sin 2\pi nx)f(x)\,dx = \frac{n}{2}b_0 \pm \frac{1}{2}b_n\,.$$

(g) Explain why the integrand above is everywhere nonnegative.

(h) Deduce that $|b_n| \leq nb_1$.

(i) Show that if $f(x) = \frac{\partial}{\partial x}P_r(x)$, then f has the required properties, and that b_n is nearly as large as nb_1 if r is near 1 and n is fixed.

6. Let $V_N(x) = 2\Delta_{2N}(x) - \Delta_N(x)$. We call this the *de la Vallée Poussin kernel*.

(a) Show that $\int_0^1 V_N(x)\,dx = 1$.

(b) Define a continuous function $v(u)$ in such a way that

$$V_N(x) = \sum_n v(n/N)e(nx)\,.$$

(c) Show that $|V_N(x)| \leq 2\Delta_{2N}(x) + \Delta_N(x)$.

(d) Show that $\int_0^1 |V_N(x)|\,dx \leq 3$.

(e) Derive an upper bound for $\int_\delta^{1-\delta} |V_N(x)|\,dx$, one that tends to 0 as N tends to infinity, for any fixed $\delta \in (0, 1/2]$.

(f) Conclude that V_N is a summability kernel.

(g) Show that the zeros of V_N are in arithmetic progression, and that every third one is a double zero.

(h) Show that $\deg V_N = 2N - 1$. (Thus by Theorem 3.15, V_N can have at most $4N - 2$ zeros.)

(i) Show that V_N has exactly $4N - 2$ zeros.

(j) Show that $-1 \leq (\sin 2\theta)^2 - (\sin\theta)^2 \leq 9/16$ for all θ. Deduce that

$$\frac{-1}{N(\sin \pi x)^2} \leq V_N(x) \leq \frac{9/16}{N(\sin\pi x)^2}$$

for all x.

7. We define the *de la Vallée Poussin power kernel* to be

$$p_n(x) = \frac{(2\cos\pi x)^{2n}}{\binom{2n}{n}}\,.$$

(a) Express $\cos \pi x$ in terms of complex exponentials, and apply the binomial theorem, to show that

$$p_n(x) = \sum_{k=-n}^{n} \frac{\binom{2n}{k+n}}{\binom{2n}{n}} e(kx).$$

(b) Show that $\int_0^1 p_n(x)\,dx = 1$.
(c) Explain why $p_n(x) \geq 0$ for all x.
(d) Deduce that $\int_0^1 |p_n(x)|\,dx = 1$ for all n.
(e) Explain why $\sum_{k=0}^{2n} \binom{2n}{k} = 2^{2n}$.
(f) Deduce that $\binom{2n}{n} \geq 2^{2n}/(2n+1)$.
(g) Deduce that $p_n(x) \leq (2n+1)(\cos \pi x)^{2n}$.
(h) Deduce that p_n is a summability kernel.
(i) What is the order of magnitude of $p_n(0)$, as a function of n? (Hint: Recall Stirling's formula, which asserts that $n! \sim \sqrt{2\pi n}(n/e)^n$.)
(j) $p_n(x)$ has a peak at $x = 0$. About how wide is this peak?

Notes

Although convergence criteria for Fourier Series were established by Dirichlet, Riemann, Dini, and Jordan, by the end of the nineteenth century it was clear convergence is problematic, and prospects for the future were guarded. Then in 1900 a twenty year old graduate student named Fejér submitted for publication a paper that revolutionized the subject. He showed that if one works with Cesàro partial sums instead of unweighted partial sums, the difficulties disappear.

For a collection of attractive proofs of the Weierstrass approximation theorem (Corollary 4.6), see Duren (2012, Chapter 6).

Recall the Dirac delta function with period 1 considered in Example 3.4. One could consider the Fejér kernel to be a Cesàro partial sum of the Fourier Series of $\delta(x)$. The lesson of Exercise 4.1.17 is that $\cot \pi x$ is the harmonic conjugate of δ. Thus the conjugate Fejér kernel is a Cesàro partial sum of the Fourier Series of $\cot \pi x$. Although $\cot \pi x \notin L^1(\mathbb{T})$, at least it is a function, whereas its conjugate, $\delta(x)$, is a distribution, but not a function. Of course in (4.53) we encountered a function $g \in L^1(\mathbb{T})$ whose conjugate $f = \widetilde{g} \notin L^1(\mathbb{T})$. M. Riesz (1928) showed that if $1 < p < \infty$ and $f \in L^p(\mathbb{T})$, then the conjugate function $\widetilde{f} \in L^p(\mathbb{T})$.

An undergraduate, Harsh Mehta, who studied from a preliminary draft of this book, asked for an REU project, and was given the task of determining the asymptotics of $\|V_N\|_1$, where V_N is the de la Vallée Poussin kernel, as defined in Exercise 4.4.6. Clearly $1 \leq \|V_N\|_1 \leq 3$, and one would expect that these numbers tend to a limit. In fact, Mehta (2014) established the surprising fact that

$$\int_0^1 |V_N(x)|\,dx = \frac{1}{3} + \frac{2\sqrt{3}}{\pi}$$

for all N.